



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA STROJNÍHO INŽENÝRSTVÍ

FACULTY OF MECHANICAL ENGINEERING

ÚSTAV PROCESNÍHO INŽENÝRSTVÍ

INSTITUTE OF PROCESS ENGINEERING

**ANALÝZA STATISTICKÉ VÝZNAMNOSTI VLIVU SOCIO-
EKONOMICKÝCH FAKTORŮ NA SLOŽENÍ ODPADU**

INFLUENCE OF SOCIO-ECONOMIC FACTORS ON THE COMPOSITION OF WASTE: STATISTICAL
APPROACH

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Petra Jordánová

VEDOUcí PRÁCE

SUPERVISOR

Ing. Radovan Šomplák, Ph.D.

BRNO 2021

Zadaní bakalářské práce

Ústav: Ústav procesního inženýrství
Studentka: **Petra Jordánová**
Studijní program: Strojírenství
Studijní obor: Základy strojního inženýrství
Vedoucí práce: **Ing. Radovan Šomplák, Ph.D.**
Akademický rok: 2020/21

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma bakalářské práce:

Analýza statistické významnosti vlivu socio–ekonomických faktorů na složení odpadu

Stručná charakteristika problematiky úkolu:

Znalost produkce a složení odpadu hraje klíčovou roli při strategickém plánování v odpadovém hospodářství. Předmětná bakalářská práce se bude zabývat vyhodnocováním vazeb mezi socio–ekonomickými daty a produkcí a složením vybraných odpadů. Student(ka) si osvojí některé statistické metody, konkrétně lze zmínit např. korelační analýzu a regresní analýzu. Nedílnou součástí bude testování statistické významnosti vytvořených modelů. Identifikované faktory dále mohou sloužit k tvorbě stratifikace území pro efektivní terénní šetření s cílem získat kvalitní odhad průměrného složení odpadu ve vybraném regionu. Bakalářská práce je motivovaná aktuálně řešeným projektem TIRSMZP719 (Prognózování produkce odpadů a stanovení složení komunálního odpadu).

Cíle bakalářské práce:

- Prohloubit si ve vhodném detailu znalosti vybraných statistických metod.
- Provést analýzu dat a vytvořit statistické modely produkce a složení odpadu.
- Testovat statistickou významnost vzniklých modelů.
- Provést diskuzi výsledků a stanovit doporučení dalšího postupu.

Seznam doporučené literatury:

BEŇO, Z. Recyklace: efektivní způsoby zpracování odpadů. Brno: Vysoké učení technické, Fakulta strojního inženýrství, Ústav procesního a ekologického inženýrství, 2011. ISBN 978-80-214-4240-5.

EDJABOU, M.E., JENSEN, M.B., GÖTZE, R., PIVNENKO, K., PETERSEN, C., SCHEUTZ, C., ASTRUP, T.F. Municipal solid waste composition: Sampling methodology, statistical analyses, and case study evaluation. Waste Management, 36, 12-23, 2015.

Termín odevzdání bakalářské práce je stanoven časovým plánem akademického roku 2020/21

V Brně, dne

L. S.

prof. Ing. Petr Stehlík, CSc., dr. h. c.
ředitel ústavu

doc. Ing. Jaroslav Katolický, Ph.D.
děkan fakulty

ABSTRAKT

Tato bakalářská práce se zabývá porovnáváním složení směsného komunálního odpadu v různých obcích České republiky. V první části je rozebrán statistický aparát potřebný pro následné vyhodnocení dat. Konkrétně je řešena náhodná veličina, rozdělení pravděpodobnosti a statistické hypotézy aj. Druhá část se zabývá analýzou dostupné datové sady. Jádrem práce je testování hypotéz o rozdílnosti složení směsného komunálního odpadu mezi vybranými obcemi a typy zástavby v obcích. V závěru je nastíněn přístup pro odhad vlivných faktorů na složení odpadu pomocí korelační a regresní analýzy. Přínosem práce je vyhodnocení stávající stratifikace pro místní šetření a doporučení změn vybraných vlivných faktorů.

ABSTRACT

This bachelor thesis deals with comparison of mixed municipal waste in different areas in the Czech Republic. Statistical apparatus needed for subsequent data is described in the first part. There is discussed problematics of random variable, the probability distribution and hypotheses that will be used in this work. The second part involves a data set analysis. The merit of this thesis is hypothesis testing of the difference in mixed municipal waste composition between chosen areas and built-up area types. At the end an approach for estimate of impact factors of waste composition using correlation and regression analysis will be given. This work brings a benefit in current stratification evaluation process and recommendation of changes in selected factors.

KLÍČOVÁ SLOVA

Směsný komunální odpad, stratifikace, složení odpadu, testování hypotéz, potenciál separace odpadu, výběrové šetření

KEYWORDS

Mixed municipal waste, stratification, waste composition, hypothesis testing, waste separation potential, sample survey

BIBLIOGRAFICKÁ CITACE

JORDÁNOVÁ, Petra. *Analýza statistické významnosti vlivu socio-ekonomických faktorů na složení odpadu* [online]. Brno, 2021 [cit. 2021-05-21]. Dostupné z: <https://www.vutbr.cz/studenti/zav-prace/detail/132857>. Bakalářská práce. Vysoké učení technické v Brně, Fakulta strojního inženýrství, Ústav procesního inženýrství. Vedoucí práce Radovan Šomplák.

PODĚKOVÁNÍ

Děkuji Ing. Radovanu Šomplákovi, Ph.D. za cenné rady a připomínky, které mi poskytl během vypracování bakalářské práce. Také děkuji za podporu projektů TIRSMZP719 Prognózování produkce odpadů a stanovení složení komunálního odpadu, TAČR a SPETEP Strategické partnerství pro environmentální technologie a produkci energie.

ČESTNÉ PROHLÁŠENÍ

Prohlašuji, že tato práce je mým původním dílem, vypracoval(a) jsem ji samostatně pod vedením vedoucího práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury.

Jako autor uvedené práce dále prohlašuji, že v souvislosti s vytvořením této práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následku porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestně právních důsledků.

V Brně dne 22. 5. 2021

.....

Petra Jordánová

OBSAH

1	ÚVOD	13
2	STATISTICKÝ APARÁT	17
2.1	Náhodná veličina	17
2.2	Číselné charakteristiky náhodné veličiny	19
2.3	Rozdělení pravděpodobnosti	20
2.4	Normální rozdělení	21
2.5	Odhady parametrů rozdělení pravděpodobnosti	24
2.5.1	Bodové odhady	24
2.5.2	Intervalové odhady	25
2.5.3	Odhady parametrů normálního rozdělení	25
2.6	Testování hypotéz	25
2.7	Korelační a regresní analýza	29
2.7.1	Typy korelačních koeficientů	29
2.7.2	Lineární regresní analýza	30
3	DATOVÁ SADA	32
4	POSTUP ZPRACOVÁNÍ DAT A PŘEDPOKLADY	35
5	VYHODNOCOVÁNÍ HYPOTÉZ	37
6	POUŽITÍ KORELAČNÍ A REGRESNÍ ANALÝZY	42
7	SHRNUTÍ A ZÁVĚRY	44
	SEZNAM POUŽITÉ LITERATURY	47
	SEZNAM POUŽITÝCH ZKRATEK	51
	SEZNAM POUŽITÝCH VELIČIN	52
	SEZNAM OBRÁZKŮ	54
	SEZNAM TABULEK	55
	SEZNAM PŘÍLOH	57

1 ÚVOD

V dnešní době je prakticky veškerá činnost člověka spojena s produkcí odpadu. Jeho tvorbou, případně jeho využitím, se zabývá oblast zvaná odpadové hospodářství (OH). Hlavním cílem je předejít produkci odpadu a pokud to není možné, najít co nejlepší řešení k jeho opětovnému využití a snížit tak spotřebu primárních surovin. Nerecyklovatelný odpad má často stále velký potenciál a může být využit na výrobu energie, čímž se opět ušetří určitá část primárních zdrojů. Poslední možností, jak lze naložit s odpadem, je jeho odstranění [1]. Nejčastějším způsobem odstranění odpadu je jeho uložení na skládku. Takovéto nakládání s odpadem není vhodné nejen z pohledu mrhání cenným materiálem, ale především z pohledu ekologické zátěže. Při skládkování je velké riziko znečištění podzemních vod. Dále při skládkování vzniká velké množství skleníkových plynů [2]. Je proto důležité, při nakládání a zpracování odpadu, myslet také na ekologický aspekt.

Každý subjekt musí v rámci OH postupovat dle platné legislativy [3] a při určité produkci odpadu jí vykazovat včetně způsobu nakládání s odpadem spádovému úřadu. Jednotlivé územní celky musí v agregované podobě plnit plány OH České republiky a podávat hodnotící zprávy o odpadech. Do plánu OH a zákonů České republiky se promítají směrnice EU. V roce 2018 byl EU schválen balíček k oběhovému hospodářství. Největší pozornost v něm je věnována komunálnímu odpadu (KO). Konkrétně se jedná o směrnice obsahující cíle pro materiálové využití KO [4], směrnice věnující se recyklaci obalů [5] a směrnice zaměřená na odklon KO od skládkování [6].

I přesto, že obce poskytují možnost třídění odpadu, jako je třeba separátní sběr papíru, plastu, skla, kovu, bio odpadu či textilu, tak stále největší část KO z hlediska hmotnosti tvoří směsný komunální odpad (SKO) [7]. Tento typ odpadu zatím není možné efektivně materiálově využít, ale s ohledem na jeho charakter je vhodný pro energetické využití [8]. Stále však obsahuje složky, které při zvýšení separace KO (přesun některých složek ze SKO do samostatně sbíraných frakcí) mohou být materiálově využity. Pro efektivní podporu zvýšení separace KO je třeba získat kvalitní odhady jeho potenciálu, tj. aktuálního složení SKO. To umožňuje stanovovat racionální cíle a plánovat potřebnou sběrnou a zpracovatelskou infrastrukturu OH v jednotlivých obcích a vyšších územních celcích.

Odhad složení SKO je většinou získán na základě jeho rozborů [9]. Složení SKO z domácností se dlouhodobě věnuje společnost EKO-KOM [10] a [11]. Dostupnost kvalitních informací o složení odpadu je obvykle omezena na nízký počet lokalit. Je to dáno technickými, časovými a finančními nároky na jednotlivé rozборы. Kvůli časové a finanční náročnosti je zapotřebí vybírat vhodné reprezentativní vzorky (viz problematika stratifikace dále v textu), jejíž

výsledky jsou následně statisticky zpracovány. Tímto způsobem je možné výsledky z vybraných lokalit agregovat pro průměrné odhady složení SKO regionů nebo států. Tato problematika zaměřena na Českou republiku je řešena v rámci aktuálně probíhajícího projektu TIRSMZP719 [12], který je realizován na Ústavu procesního inženýrství (ÚPI), Vysoké učení technického v Brně. Plánované řešení projektu je 2019 až 2021. S ohledem na omezené terénní práce v rámci okolností vyplývajících z pandemie Covid-19 bude možná řešení projektu prodlouženo. S tím souvisí omezené množství dat z rozborů odpadu v době psaní této bakalářské práce. Z tohoto důvodu bylo nutné vhodně upravit použité metody pro vyhodnocení dílčích výsledků. V některých případech byly přístupy pouze diskutovány na teoretické úrovni a praktické použití bude možné realizovat až při rozšíření datové sady po obnovení terénních prací v rámci projektu TIRSMZP719.

Ke zjištění složení SKO je nutné zvážit použití stratifikace. Především se jedná o situaci, kdy je cílem odhadnout složení pro celé regiony nebo státy. Aby byla stratifikace co nejúčinnější, je třeba zvolit vhodné faktory, které mají zásadní vliv na sledované parametry, tedy na složení SKO. ÚPI navázal spolupráci v rámci projektu SPETEP [13] s Univerzitou Karlovou v Praze, kde se řešitelé společně s odborníky z Matematicko-fyzikální fakulty zabývali problematikou plánování a statistického vyhodnocování výběrových šetření (rozborů složení odpadů). Na základě odborné debaty v rámci řešitelských týmů projektů TIRSMZP719 a SPETEP byly expertně navrženy vhodné socio-ekonomické a jiné parametry a na jejich základě byla provedena následná stratifikace. Zvoleným přístupem pro stratifikaci bylo vybráno 10 strat vždy o jednom reprezentantovi. Konkrétní metodou použité stratifikace se zabývá odborná studie, která vznikla v rámci projektu SPETEP: *Vývoj statistické metodiky pro místní šetření – rozborů složení směsného komunálního odpadu na úrovni producentů odpadu*. Výsledky výzkumu byly promítnuty do dokumentu projektu TIRSMZP719: *Certifikovaná metodika pro stanovení složení směsného komunálního odpadu z obcí a komunálního odpadu*. Po stratifikaci začalo pro vybrané reprezentanty šetření zaměřené na rozborů složení SKO, které však bylo vlivem pandemie pozastaveno, a jsou tedy dostupná, ve chvíli psaní tohoto textu, jen data z obcí reprezentující tři velikosti – krajské město, město o velikosti obce s rozšířenou působností a obce vesnického typu.

V ideálním případě by se vybraly faktory (např. socio-ekonomická data) pro stratifikaci na základě dat o složení SKO z potřebného množství obcí. Tato sada není v České republice v současnosti k dispozici, proto bylo přistoupeno k výběru parametrů stratifikace na základě expertního odhadu (viz výše). V dalším kroku je tedy nutné vhodnost expertního výběru faktorů ověřit na nově vznikající datové sadě. Kvůli pandemii není v tuto chvíli k dispozici potřebná datová sada, i přesto, je ale již v tuto chvíli možné učinit dílčí testování a připravit teoretický výpočtový aparát.

Tato bakalářská práce navrhuje přístup, jakým bude možné výsledky stratifikace opakovaně hodnotit při získávání nových datových sad z rozborů odpadu. Z výsledků bude možné seznam reprezentativních obcí pro vykonání rozborů SKO upravit a naplánovat nové lokality pro rozborů, aby byl odhad průměrného složení SKO v České republice dostatečně přesný při minimálním (ve smyslu aktuálně dostupných dat) počtu rozborů.

Vyhodnocování datových sad je možné rozdělit na dvě části. První část se týká stratifikace na úrovni obcí. Na této úrovni je uvažováno 6258 obcí na území České republiky. Zde bylo provedeno dílčí (s ohledem na data pouze ze tří obcí) testování hypotéz s vazbou na kvalitu vybraných reprezentantů. Dále v této části je vhodné využít korelační a regresní analýzu mezi vybranými faktory a výsledky pro jednotlivá strata (obce) pro ověření správnosti vybraných parametrů stratifikace. Tato analýza je řešena pouze teoreticky s ohledem na množství dostupných dat. Ve druhé části jde o stratifikaci na místní úrovni (výběr sběrných nádob), kde hrají vliv například: typ zástavby, ústřední vytápění a další faktory týkající se přímo jednotlivých obcí. V této oblasti proběhlo opět testování hypotéz o správnosti stratifikace na úrovni sběrných nádob ve vybraných obcích.

2 STATISTICKÝ APARÁT

Nejprve, než se začne aplikovat statistický aparát na reálna data je vhodné uvést pár základních pojmů. Pro výpočet pravděpodobnosti je z praktických důvodů vhodné k jednotlivým jevům základního souboru (výsledkům jednotlivých pokusů) přiřadit číselné hodnoty. Toto přiřazení určuje funkce zvaná náhodná veličina. Bude zde také pojednáno o jejích číselných charakteristikách. Více podrobností a konkrétních definic je možno nalézt v [14] a [15].

Další část této kapitoly se bude týkat rozdělení pravděpodobnosti, které se bude věnovat především normálnímu rozdělení [14]. Budou zde také popsány bodové a intervalové odhady pro jeho hlavní charakteristiky [15]. Dále budou popsány konkrétní statistické testy, které budou následně využity v aplikační části bakalářské práce [15] a [16].

Bude diskutována problematika statistických hypotéz, na jejíž základě je možné stanovit následný postup pro vyhodnocování dalších datových sad.

Poslední část této kapitoly bude věnována korelační a lineární regresní analýze.

2.1 Náhodná veličina

Náhodná veličina je značena velkým písmenem X . Je to proměnná, která náhodně nabývá číselných hodnot x , které jsou získány realizací náhodných pozorování. Náhodná veličina může být přesněji definována jako funkce, která přiřazuje každému elementárnímu náhodnému jevu nějakou hodnotu. Náhodná veličina může být představena jako počet ok při házení kostkou na obr. 1 a na obr. 2, případně např. měření teploty ve stejné denní době. Náhodná veličina se dělí na dva typy [15].

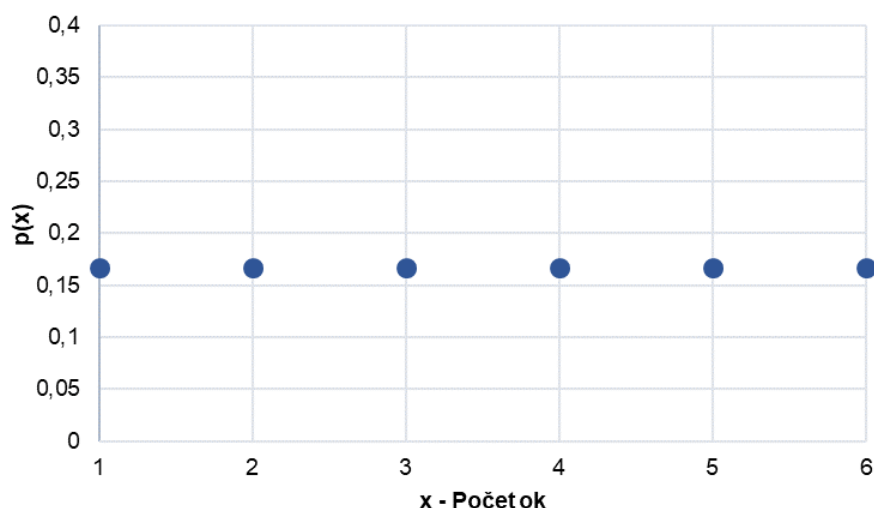
Prvním uvedeným typem je náhodná veličina diskrétního (nespojitého) typu. Diskrétní náhodná veličina může nabývat nejvýše spočetně mnoha hodnot z konečné i nekonečné množiny. Je popsána pravděpodobností funkcí a její sumace se musí rovnat jedné.

$$p(x) = P(X = x); p(x) \geq 0 \text{ pro } x = x_1, x_2, \dots \quad (1)$$

Dále je popsána distribuční funkcí $F(x)$, která má schodovitý tvar a říká, s jakou pravděpodobností se náhodná veličina X vyskytne v daném intervalu (souboru hodnot), tj. (dle [15], [16] a [17]).

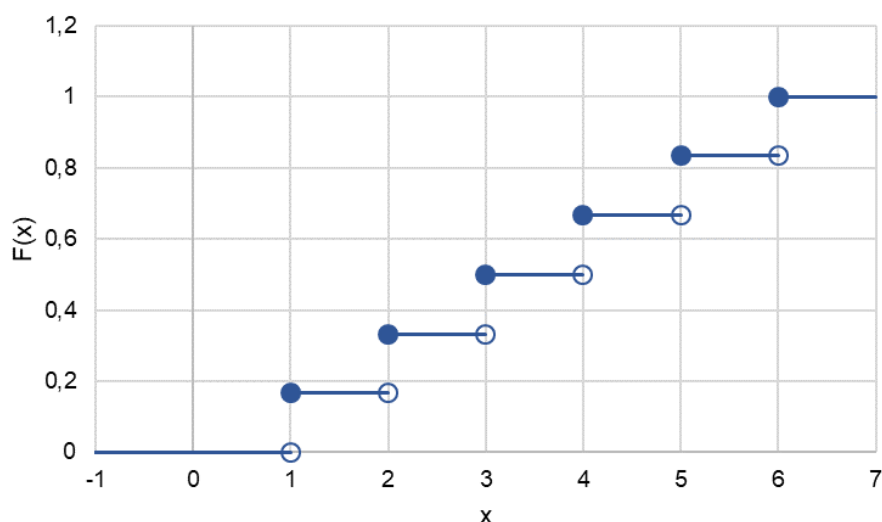
$$F(x) = P(X \leq x). \quad (2)$$

Na obr. 1 je zobrazena pravděpodobnostní funkce pro hod kostkou.



Obr. 1 Příklad grafu pravděpodobnostní funkce diskrétního typu

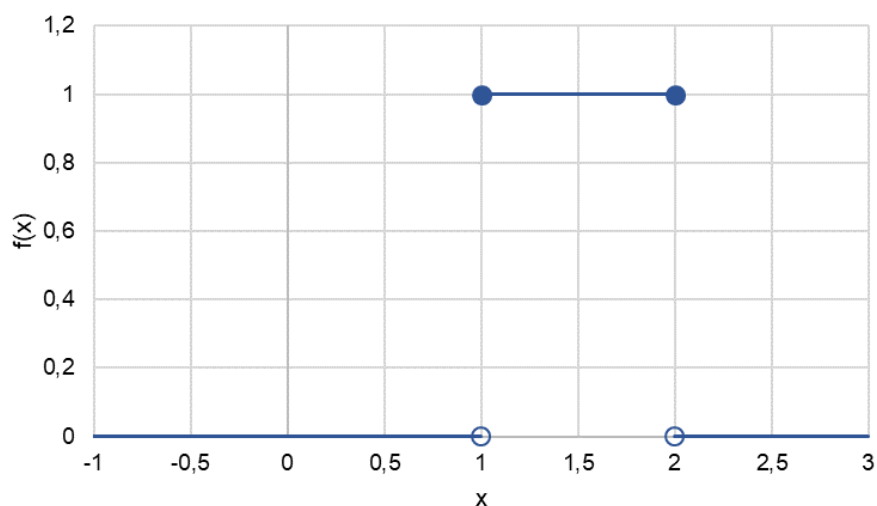
Na obr. 2 je zobrazena distribuční funkce pro hod kostkou.



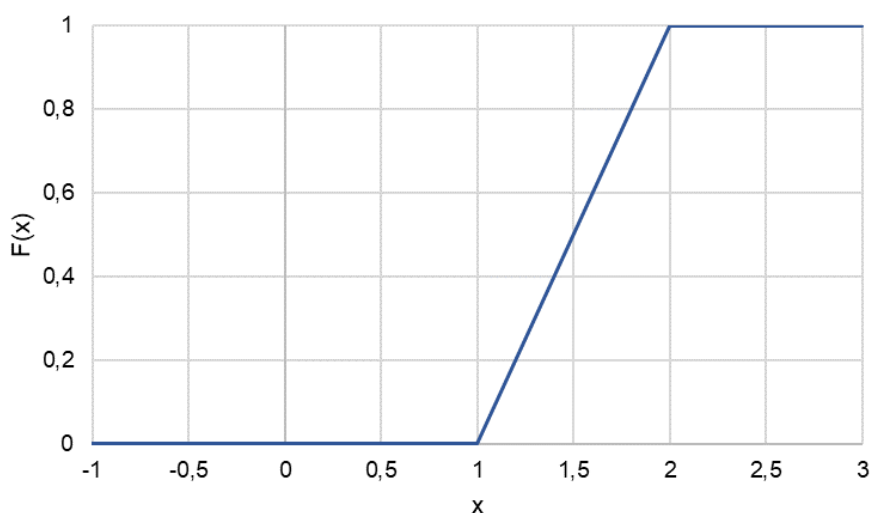
Obr. 2 Příklad grafu distribuční funkce diskrétního typu

Druhým a pro tuto práci důležitějším typem je náhodná veličina spojitého typu (viz [15] a [17]). Spojitá náhodná veličina může nabývat všech hodnot z konečného i nekonečného intervalu. Spojitá náhodná veličina je popsána hustotou pravděpodobnosti $f(x)$ (viz obr. 3), která udává, s jakou četností se zkoumaný jev vyskytne v celém souboru zkoumaných hodnot. Je to nezáporná funkce definovaná na intervalu $(-\infty, \infty)$. Stejně jako diskrétní náhodná veličina je i ta spojitá popsána distribuční funkcí (viz obr. 4), která už však není schodovitého tvaru, ale je spojitá na celé délce intervalu a udává, s jakou pravděpodobností se zkoumaná veličina vyskytne ve zkoumaném souboru hodnot. Hustota pravděpodobnosti a distribuční funkce jsou svázány vztahem

$$F(x) = \int_{-\infty}^x f(t)dt \text{ a platí } \int_{-\infty}^{+\infty} f(x)dx = 1. \quad (3)$$



Obr. 3 Příklad grafu hustoty pravděpodobnosti spojitého typu



Obr. 4 Příklad grafu distribuční funkce spojitého typu

2.2 Číselné charakteristiky náhodné veličiny

Číselné charakteristiky jsou reálná čísla, která vyjadřují důležité vlastnosti náhodné veličiny X . První uvedenou číselnou charakteristikou je střední hodnota $E(X)$. Ta udává polohu rozdělení pravděpodobnosti náhodné veličiny X . Střední hodnota se dá interpretovat také fyzikálně například při zjišťování těžiště tělesa. Pro diskrétní náhodnou veličinu je popsána váženým průměrem všech jejích možných hodnot:

$$E(X) = \sum_x xp(x). \quad (4)$$

Pro spojitou náhodnou veličinu je dána integrálem proměnné vzhledem k její hustotě pravděpodobnosti $f(x)$:

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx. \quad (5)$$

Další číselnou charakteristikou je rozptyl neboli střední kvadratická odchylka nebo také disperze či variabilita. Udává, jak moc se jednotlivé hodnoty náhodné veličiny X liší od její střední hodnoty. V jednotlivých literaturách je značena rozdílnými způsoby: $\sigma^2(x)$ nebo $D(X)$. Střední kvadratická odchylka může být interpretována také fyzikálně, a to jako moment setrvačnosti. Tato číselná charakteristika nabývá vždy nezáporných hodnot a je definována následujícím vztahem:

$$D(X) = E([X - E(X)]^2). \quad (6)$$

Pro diskrétní náhodnou veličinu X platí

$$D(X) = \sum_x (x - E(X))^2 p(x) = \sum_x x^2 p(x) - (E(X))^2. \quad (7)$$

Pro spojitou náhodnou veličinu X platí

$$D(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx = \int_{-\infty}^{+\infty} x^2 f(x) dx - (E(X))^2. \quad (8)$$

Poslední důležitou číselnou charakteristikou, která zde je uvedena, je směrodatná odchylka $\sigma(X)$, která je definována jako odmocnina z rozptylu náhodné veličiny X . Stejně jako rozptyl i směrodatná odchylka nabývá pouze nezáporných hodnot a čím větší je směrodatná odchylka, tím větší je i rozptyl. Směrodatná odchylka udává, jak moc se hodnoty liší od střední hodnoty. Směrodatná odchylka je užívána i v jiných oborech, než je matematika, například v metrologii při kontrole rozměrů součástí. Definiční vztah pro diskrétní i spojitou náhodnou veličinu je následující:

$$\sigma(X) = \sqrt{D(X)}. \quad (9)$$

Číselné charakteristiky náhodné veličiny byly převzaty z [15] a [16].

2.3 Rozdělení pravděpodobnosti

Rozdělení pravděpodobnosti náhodné veličiny přiřadí náhodnému jevu, který je popsán náhodnou veličinou, určitou pravděpodobnost. Tato pravděpodobnost může být diskrétního či spojitého typu. Jednotlivá rozdělení pravděpodobnosti se liší především tvarem hustoty pravděpodobnostní funkce, viz kapitola 2.1 *Náhodná veličina*. Tvar křivky je dán vztahem

pro jednotlivé rozdělení a charakteristickými parametry rozdělení. Každé rozdělení pravděpodobnosti má své určité vlastnosti, podle kterých lze určit, jaké rozdělení je nejvíce vhodné pro popis sledovaného jevu.

Prvním uvedeným typem je rozdělení pravděpodobnosti diskrétní náhodné veličiny, které přiřazuje náhodnému pokusu právě jednu hodnu. Toto rozdělení může být dále děleno na jednotlivá rozdělení, která mají své specifické vlastnosti, a využívají se u konkrétních úloh. Každé rozdělení je popsáno svými vzorci a má své vlastní značení. K důležitým rozdělením diskrétního typu například patří: Alternativní (Bernoulliho) rozdělení $A(p)$, Klasické rozdělení $C(n)$, Binomické rozdělení $Bi(n, p)$, Hypergeometrické rozdělení $H(N, M, n)$, Poissonovo rozdělení $Po(\lambda)$ a Geometrické rozdělení $Ge(p)$. Více informací o diskrétním rozdělení pravděpodobnosti je k nalezení v [14] a [15].

Druhým typem je rozdělení pravděpodobnosti spojité náhodné veličiny. Toto rozdělení nelze popsat pravděpodobnostní funkcí v určitém bodě, a proto byla zavedena hustota pravděpodobnosti, viz kapitola 2.1 *Náhodná veličina* a její distribuční funkce je spojitá. I toto rozdělení lze dále dělit na jednotlivá rozdělení spojitého typu, která mají své značení a vzorce. K důležitým rozdělením spojitého typu patří například: Rovnoměrné rozdělení $R(a, b)$, Exponenciální $Exp(A, \delta)$, Weibullovo rozdělení $W(\delta, b, t)$, Pearsonovo χ^2 rozdělení $\chi^2(k)$ a Studentovo t rozdělení $S(k)$. Posledním a pro tuto práci nejdůležitějším uvedeným rozdělení pravděpodobností spojité náhodné veličiny je Normální (Gaussovo) rozdělení, které je podrobněji rozebráno v následující kapitole. Více informací o spojitém rozdělení pravděpodobnosti je k nalezení v [14] a [15].

2.4 Normální rozdělení

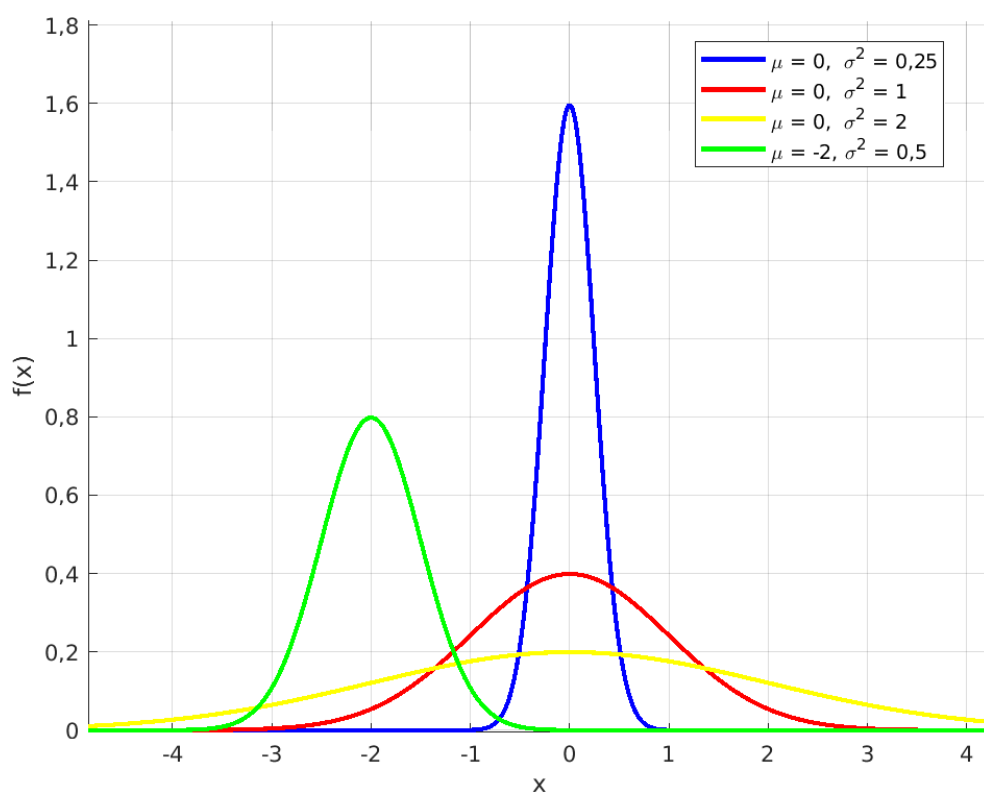
Normální, též Gaussovo, rozdělení je nejdůležitější rozdělení pravděpodobnosti spojité náhodné veličiny. Oproti jiným rozdělením má toto rozdělení velký význam pro teoretickou statistiku (např. Centrální limitní věta, která říká, že pokud náhodná veličina X vznikne součtem velkého počtu vzájemně nezávislých náhodných veličin z libovolného, ale stejného rozdělení, pak bude mít normální rozdělení) [18]. Normální rozdělení je znázorněno tzv. Gaussovou křivkou, která svým typickým zvonovitým tvarem popisuje hodnoty veličin, shlukujících se kolem střední hodnoty. Náhodné pokusy jsou často ovlivněny velkým počtem vzájemně slabě působících faktorů, které mají ve výsledku přibližně normální rozdělení. Příkladem může být např. výška člověka, maximální dosažená rychlost zvířete či člověka na určitou vzdálenost, ale i kupříkladu průměrné složení sběrných nádob pro odpad.

Normální rozdělení je značeno $N(\mu, \sigma^2)$, kde $\mu, \sigma^2 \in R, \sigma^2 > 0$. Parametr μ představuje střední hodnotu $E(X)$ normálního rozdělení pravděpodobnosti a jeho změněním dochází k posunu středu Gaussovy křivky na ose x . Druhý parametrem je σ^2 , které představuje rozptyl $D(X)$ normálního rozdělení. Při jeho změně dochází k zúžení či rozšíření Gaussovy křivky, viz obr. 5 a obr. 6. Čím je rozptyl menší, tím je křivka užší, a tedy i vyšší a naopak.

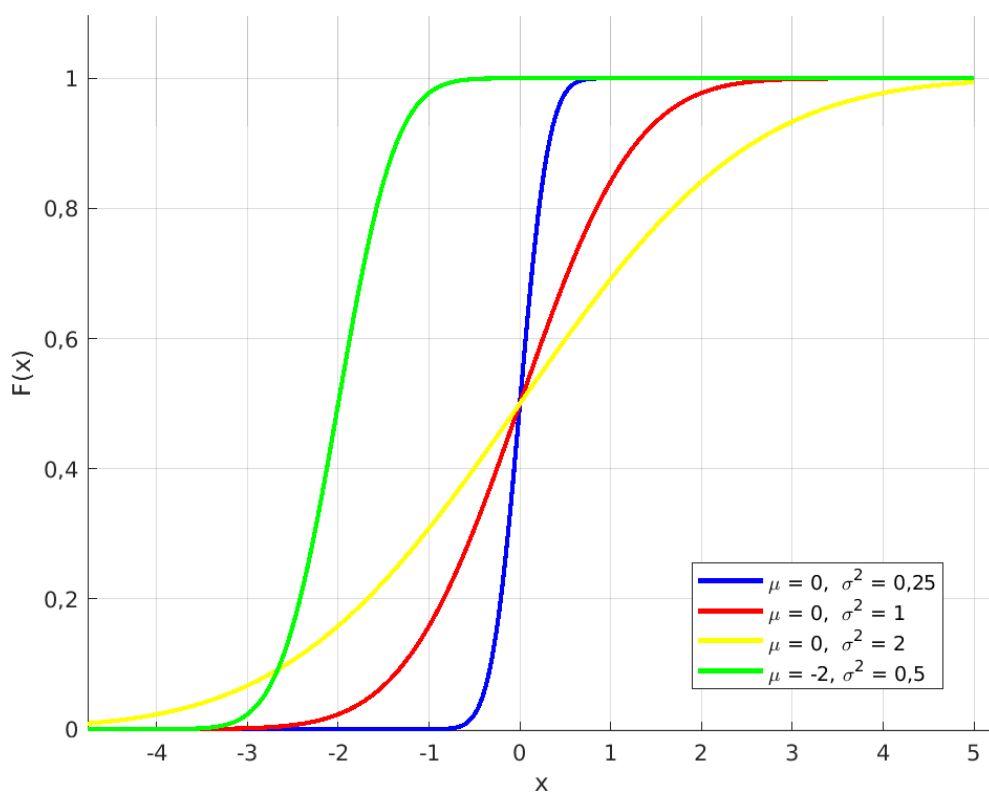
Normální rozdělení lze popsat vzorcem pro hustotu pravděpodobnosti

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \text{ kde } x \in R. \quad (10)$$

Normální rozdělení převzato z [15] a [19].



Obr. 5 Graf hustoty normálního rozdělení



Obr. 6 Graf distribuční funkce normálního rozdělení

2.5 Odhady parametrů rozdělení pravděpodobnosti

Číselné charakteristiky často nelze přesně určit, je tedy třeba je alespoň odhadnout, a proto jsou zavedeny bodové a intervalové odhady parametrů. Než bude pojednáno o rozdělení odhadů, je vysvětlen pojem parametru ϑ (viz [15] a [20]).

Parametr ϑ je reálné číslo nebo reálný vektor rozdělení pravděpodobnosti. Obvykle je hodnota parametru neznámá a je odhadována na základě získaného statistického souboru (x_1, \dots, x_n) pro náhodnou veličinu nebo případně $((x_1, y_1), \dots, (x_n, y_n))$ pro náhodný vektor. Pro následující výpočty bude parametr ϑ používán jako číselné charakteristiky náhodné veličiny (náhodného vektoru), viz kapitola 2.1 *Náhodná veličina*.

2.5.1 Bodové odhady

Bodový odhad je pozorovaná hodnota $t = T(x_1, \dots, x_n)$ ze statistického souboru (x_1, \dots, x_n) . U bodových odhadů jsou používány především odhady:

1. **Nestranný (nevychýlený)**, pokud střední hodnota odhadu T je rovna parametru ϑ neboli

$$E(T) = \vartheta. \quad (11)$$

2. **Nejlepší nestranný odhad** nastane, pokud rozptyl nestranného odhadu T je nejmenší z rozptylů všech nestranných odhadů téhož parametru ϑ .
3. Odhad T je **konzistentní**, platí-li

$$\lim_{n \rightarrow \infty} P(|T - \vartheta| < \varepsilon) = 1 \quad (12)$$

pro libovolné reálné číslo $\varepsilon > 0$.

U bodového odhadu pro číselné charakteristiky náhodné veličiny X platí, že: \bar{x} je bodovým odhadem střední hodnoty $E(X)$, s^2 je bodovým odhadem rozptylu $D(X)$ a s je bodovým odhadem směrodatné odchylky $\sigma(X)$, kde \bar{x} , s^2 a s jsou empirické charakteristiky získané ze statického souboru (x_1, \dots, x_n) a jsou popsány následujícími vztahy (viz [16] a [21]):

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (13)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (14)$$

$$s = \sqrt{s^2}. \quad (15)$$

2.5.2 Intervalové odhady

U intervalových odhadů je třeba zavést interval spolehlivosti, který pro parametr ϑ se spolehlivostí $1 - \alpha$, kde $\alpha \in (0; 1)$, je dvojice T_1 a T_2 takových statistik, že $P(T_1\vartheta \leq T_2) = 1 - \alpha$ pro libovolnou hodnotu parametru ϑ .

Intervalový odhad T parametru ϑ říká, s jakou pravděpodobností leží jeho hodnota v aproximovaném intervalu. Pro intervalový odhad parametru ϑ se spolehlivostí $1 - \alpha$ platí $\vartheta \in \langle t_1; t_2 \rangle$, kde t_1, t_2 jsou hodnoty statistik T_1, T_2 statistického souboru (x_1, \dots, x_n) .

Spolehlivost $1 - \alpha$ je volena blízká jedné, obvykle 0,95 nebo 0,99 a uvádí se v procentech. Spolehlivost $1 - \alpha$ říká, kolikrát (z kolika procent – $(1 - \alpha)100\%$) bude hodnota parametru ϑ obsažena ve zvolených intervalových odhadech z daného základního souboru při opakovaných výběrech s konstantním rozsahem n [16] a [22].

2.5.3 Odhady parametrů normálního rozdělení

Normální rozdělení představuje významné rozdělení, jak bylo již uvedeno v kapitole 2.4 *Normální rozdělení*. Toto rozdělení je často využíváno pro popis technických parametrů, které je třeba odhadnout. Vzorce pro bodové odhady \bar{x} , s^2 a s jsou uvedeny v kapitole 2.5.1 *Bodové odhady* rovnice (13), (14) a (15).

Intervalový odhad střední hodnoty μ se spolehlivostí $1 - \alpha$, při neznámém rozptylu σ^2 je

$$\langle \bar{x} - t_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}; \bar{x} + t_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \rangle,$$

kde $t_{1-\frac{\alpha}{2}}$ je $(1 - \frac{\alpha}{2})$ kvantil Studentova rozdělení $S(k)$ s $k = n - 1$ stupni volnosti.

Intervalový odhad rozptylu σ^2 se spolehlivostí $1 - \alpha$ je

$$\langle \frac{ns^2}{\chi^2_{1-\frac{\alpha}{2}}}; \frac{ns^2}{\chi^2_{\frac{\alpha}{2}}} \rangle,$$

kde χ^2_P je P – kvantil Pearsonova rozdělení $\chi^2(k)$ s $k = n - 1$ stupni volnosti [16] a [23].

2.6 Testování hypotéz

Testování hypotéz slouží k určení pravdivosti tvrzení neboli statistické hypotézy o parametrech náhodné veličiny. Test statistické hypotézy je popsán matematickými postupy, které vedou k zamítnutí či nezamítnutí dané hypotézy. Statistická hypotéza H popisuje vlastnosti rozdělení pravděpodobnosti pozorované náhodné veličiny X s distribuční funkcí $F(x, \vartheta)$. Hypotéza může být nulová či alternativní, ta může být dále dělena na jednostrannou a dvoustrannou. Pro nulovou hypotézu platí, že $H: \vartheta = \vartheta_0$, tedy výsledkem je právě jedna

hodnota. Pro jednostrannou hypotézu platí: $H: \vartheta > \vartheta_0$ resp. $H: \vartheta < \vartheta_0$, tedy výsledkem je interval hodnot, který je větší či menší než zvolená hodnota nulové hypotézy. Pro dvoustrannou alternativní hypotézu platí: $\bar{H}: \vartheta \neq \vartheta_0$, výsledkem jsou tedy všechny hodnoty mimo hodnotu nulové hypotézy. Obdobně je to i pro alternativní hypotézy u jednostranné varianty.

Pro potvrzení či zamítnutí hypotézy je třeba také zmínit kritický obor W_α a jeho doplněk \bar{W}_α . Index α je hladina významnosti testu a obvykle se volí 0,05 nebo 0,1, může být vyjádřena v procentech.

Posledním uvedeným pojmem je pozorovaná hodnota testového kritéria, která je definována $t = T(x_1, \dots, x_n)$, kde $T(x_1, \dots, x_n)$ je testové kritérium.

Závěrečné možnosti vyhodnocení hypotézy je uvedeno v tab. 1.

Tab. 1 Tabulka závěrečného vyhodnocení hypotéz

	Skutečnost		
	H	Platí	Neplatí
Rozhodnutí	Zamítáme	Chyba 1. druhu	Vyhovující fakt
	Nezamítáme	Vyhovující fakt	Chyba 2. druhu

Úvod kapitoly testování hypotéz byl převzat z [16].

U datových sad je obvykle cílem zjistit, zda mají podobné střední hodnoty a rozptyly. Testy na porovnání vlastností datových sad mohou být parametrické (mají vyšší míru testu, tzn. menší chybovost druhého druhu) či neparametrické. Pro neparametrické hypotézy není nutná znalost typu rozdělení datových sad. U parametrických hypotéz je znalost typu rozdělení nutná, a proto bude jako první uveden Kolmogorovův-Smirnovův test, který testuje normalitu dat. Při nezamítnutí testu normality dat budou následovat testy pro porovnávání středních hodnot a následně rozptylů.

Kolmogorovův-Smirnovův test

Kolmogorovův-Smirnovův test slouží k ověření, zda soubor získaných náhodných veličin pochází z normálního rozdělení.

Testování hypotézy je založeno na porovnávání výběrové distribuční funkce $F_n(x)$ s hypotetickou distribuční funkcí $F_0(x)$. Testovou statistickou je

$$D_n = \max_{1 \leq x \leq n} |F_n(x) - F_0(x)|. \quad (16)$$

Nulovou hypotézu je zamítnuta na hladině významnosti α , když $D_n \geq D_n(\alpha)$, kde $D_n(\alpha)$ je tabelovaná kritická hodnota. Hodnoty $D_n(\alpha)$ jsou uvedeny v tabulce, je v přílohách (viz *Příloha 1*) [24] a [25].

Než budou použity testy pro porovnávání středních hodnot, musí být nejprve proveden test na porovnávání rozptylů. Test pro porovnání středních

hodnot může být jednovýběrový, párový a dvouvýběrový při stejných či různých rozptylech. V této práci jsou použity pouze dvouvýběrové testy pro porovnání středních hodnot. Více o párovém a jednovýběrovém testu je možné zjistit více z [15]. Test pro dva výběry při stejných rozptylech bude použit, pokud nebude zamítnut test pro porovnávání rovnosti rozptylů. Jestliže test pro porovnání rozptylů bude zamítnut, poté bude použit test pro dva výběry při různých rozptylech.

F-test (Fisherův test)

Test hypotézy $H: \sigma^2(X) = \sigma^2(Y)$. Fisherův test určí, zda hodnoty dvou porovnávaných souborů mají stejné rozptyly. Pozorovaná hodnota testovaného kritéria je

$$t = \frac{\max\left(\frac{n_1 s^2(x)}{n_1 - 1}, \frac{n_2 s^2(y)}{n_2 - 1}\right)}{\min\left(\frac{n_1 s^2(x)}{n_1 - 1}, \frac{n_2 s^2(y)}{n_2 - 1}\right)}, \quad (17)$$

kde klademe $\bar{W}_\alpha \in \langle 1; F_{1-\frac{\alpha}{2}} \rangle$ a $F_{1-\frac{\alpha}{2}}$ je $(1 - \frac{\alpha}{2})$ – kvantil Fisherova-Snedecorova rozdělení $F(k_1, k_2)$ se stupni volnosti $k_1 = n_1 - 1$ a $k_2 = n_2 - 1$ pro $\frac{n_1 s^2(x)}{n_1 - 1} \geq \frac{n_2 s^2(y)}{n_2 - 1}$ anebo $k_1 = n_2 - 1$ a $k_2 = n_1 - 1$ pro $\frac{n_1 s^2(x)}{n_1 - 1} \leq \frac{n_2 s^2(y)}{n_2 - 1}$. Kvantily pro toto rozdělení nalezneme v tabulkách (*Příloha 2* a *Příloha 3*), viz [15] a [16].

t-test (Studentův test pro dva výběry při stejných rozptylech)

Test hypotézy $H: \mu(X) - \mu(Y) = \mu_0$ při neznámých rozptylech $\sigma^2(X) = \sigma^2(Y)$. Je určen pro porovnání střední hodnoty dvou skupin. Pozorovaná hodnota testovaného kritéria je

$$t = \frac{\bar{x} - \bar{y} - \mu_0}{\sqrt{\frac{n_1 s^2(x) + n_2 s^2(y)}{n_1 + n_2}}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \quad (18)$$

a $\bar{W}_\alpha = \langle -t_{1-\frac{\alpha}{2}}, t_{1-\frac{\alpha}{2}} \rangle$, kde $t_{1-\frac{\alpha}{2}}$ je $(1 - \frac{\alpha}{2})$ – kvantil Studentova rozdělení $S(k)$ s $k = n_1 + n_2 - 2$ stupni volnosti, hodnoty jsou uvedeny v tabulkách (*Příloha 4*) [15] a [25].

t-test (Studentův test pro dva výběry při různých rozptylech)

Test hypotézy $H: \mu(X) - \mu(Y) = \mu_0$ při neznámých rozptylech, tedy $\sigma^2(X) \neq \sigma^2(Y)$. Je používán v případech jako u předešlého testu, pouze pro výběry s různými rozptily. Pozorovaná hodnota testovaného kritéria je

$$t = \frac{\bar{x} - \bar{y} - \mu_0}{\sqrt{\frac{s^2(x)}{n_1 - 1} + \frac{s^2(y)}{n_2 - 1}}} \quad (19)$$

a $\bar{W}_\alpha = \langle -\bar{t}_{1-\frac{\alpha}{2}}; \bar{t}_{1-\frac{\alpha}{2}} \rangle$, kde $t_{1-\frac{\alpha}{2}} = \frac{\frac{s^2(x)}{n_1-1}t(x) + \frac{s^2(y)}{n_2-1}t(y)}{\frac{s^2(x)}{n_1-1} + \frac{s^2(y)}{n_2-1}}$ a $t(x), t(y)$ je $(1 - \frac{\alpha}{2})$ – kvantil

Studentova rozdělení $S(k)$ s $k = n_1 - 1$, resp. $k = n_2 - 1$, stupni volnosti, hodnoty jsou uvedeny v tabulkách (*Příloha 4*) [15] a [25].

Jestliže byl zamítnut test, který zkoumal normalitu dat, nemohou být výše zmíněné testy použity. Pro datové sady, které nepocházejí z normálního rozdělení, je možné použít neparametrický dvouvýběrový Wilcoxonův test. Tímto testem bude zjištěno, zda dva soubory mají podobné vlastnosti.

Dvouvýběrový Wilcoxonův test

Dvouvýběrový Wilcoxonův test se používá pro dva výběry X_1, \dots, X_2 a Y_1, \dots, Y_2 o neznámém rozdělení. Zkoumá se, zda rozdělení v souboru dat X je shodné H či odlišné \bar{H} od rozdělení v souboru dat Y . Tento test se řadí mezi pořadové testy. Na rozdíl od t-testů, které porovnávají střední hodnoty, pracuje Wilcoxonův test s celým souborem dat obou výběrů.

Test pracuje s výběry X a Y jako s celkem. V prvním kroku seřadí všechna získaná data vzestupně a přiřadí jim hodnoty pořadí $1 - n$. Aby platila nulová hypotéza, musí mít daný znak v obou populacích stejné rozdělení, tedy hodnoty přiřazené datům by měli být dobře promíchány.

Pro výpočet je třeba udělat součet pořadí od jednoho z výběrů např. Y_1, \dots, Y_2

$$W_y = \sum_{i=1}^{n_y} R(Y_i), \quad (20)$$

kde $R(Y_i)$ značí pořadí Y_i . Není známé přesné rozdělení veličiny W_y , a proto bude použita aproximace normálním rozdělením, kde n_x a n_y značí rozsah výběrů náhodných veličin v datovém souboru. Pro tento test je použit vztah

$$Z = \frac{W_y - \frac{1}{2}n_y(n_x + n_y + 1)}{\sqrt{n_x n_y (n_x + n_y + 1) \frac{1}{12}}}. \quad (21)$$

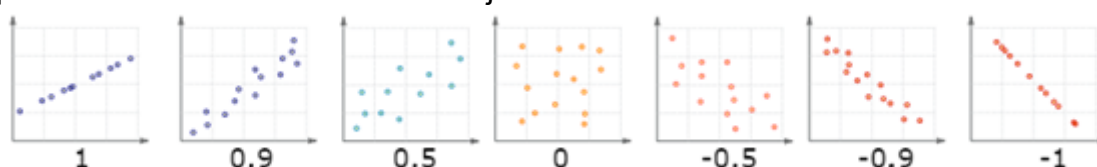
Hypotéza o shodě rozdělení je zamítnuta, jestliže je hodnota Z příliš velká. Hypotézu H tedy zamítneme, pokud platí

$$|Z| \geq qnorm\left(1 - \frac{\alpha}{2}\right), \quad (22)$$

kde $qnorm\left(1 - \frac{\alpha}{2}\right)$ je hodnota, která má stření hodnotu rovnu nule a směrodatnou odchylku rovnou jedné. Jedná se inverzní funkci k normálnímu rozdělení [27].

2.7 Korelační a regresní analýza

Korelační analýza slouží k vyhodnocování vzájemného vztahu dvou a více náhodných veličin. Korelační analýza slouží nejen k určení vzájemného vztahu náhodných veličin, ale také k určení, jak moc je tato závislost silná. Sílu závislosti popisují korelační koeficienty. Hodnota korelačního koeficientu r se pohybuje v intervalu $\langle -1, 1 \rangle$, kde znaménko určuje směr korelace. Pro koeficienty platí, že čím vyšší je hodnota koeficientu v absolutní hodnotě, tím silnější je vztah mezi náhodnými veličinami [28]. Ukázka výsledných hodnot korelačních koeficientů pro Pearsonův koeficient korelace je znázorněna na obr. 7.



Obr. 7 Příklad hodnot korelačních koeficientů [29]

2.7.1 Typy korelačních koeficientů

Pearsonův korelační koeficient

Pearsonův korelační koeficient $R(X, Y)$ je využíváný pouze pro lineární závislosti. Pro vyhodnocování nelineárních vztahů není vhodný. Aby mohl být Pearsonův korelační koeficient počítán musí data pocházet z normálního rozdělení. A rozsah výběru nesmí být příliš malý. Hodnoty tohoto koeficientu se pohybují na intervalu $\langle -1, 1 \rangle$. Jestliže vyšší hodnota náhodné veličiny X souvisí s vyšší hodnotou náhodné veličiny Y či naopak, pak je hodnota $R(X, Y)$ kladná. Pearsonův korelační koeficient je vyjádřen následujícím vztahem

$$R(X, Y) = \frac{E((X - EX)(Y - EY))}{\sqrt{DX}\sqrt{DY}}. \quad (23)$$

Pro výpočet Pearsonova korelačního koeficientu je možné použít i následující vzorec

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{(n-1)s_x s_y}. \quad (24)$$

Pearsonův korelační koeficient byl převzat z [28].

Spearmanův korelační koeficient

Spearmanův korelační koeficient r_s je neparametrický koeficient korelace, který slouží k hodnocení nelineárních závislostí (hodnotí monotónní vztah). Pro výpočet tohoto koeficientu je použita hodnota x_{ri} jako pořadí hodnoty x_i v rámci vzestupně uspořádaného souboru x_1, \dots, x_n a hodnota y_{ri} jako pořadí hodnoty y_i v rámci vzestupně uspořádaného souboru y_1, \dots, y_n , dále hodnota \bar{x}_r, \bar{y}_r jako jejich průměry a s_{xr} a s_{yr} jako jejich směrodatné odchylky. Pro výpočet Spearmanova korelačního koeficientu je použitý vzorec

$$r_s = \frac{\sum_{i=1}^n x_{ri}y_{ri} - n\bar{x}_r\bar{y}_r}{(n-1)s_{xr}s_{yr}}. \quad (25)$$

Spearmanův korelační koeficient může být také počítán pomocí diferencí získaných hodnot. Poté je možné vztah vyjádřit jako

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}, \quad (26)$$

kde platí, že $d_i = x_{ri} - y_{ri}$ [28].

2.7.2 Lineární regresní analýza

Regresní analýza je využívána pro popis vazeb mezi vybranými veličinami. Je možné ji využít po provedení dostatečného množství experimentů. Jako příklad může být uvedeno určení průměrného složení SKO, pomocí vysvětlujících faktorů spojených s jednotlivými obcemi, kde byl rozbor odpadu proveden. Lineární regrese lze zapsat následující rovnicí.

$$Y_i = \beta_0 + \sum_{i \in I} (\beta_i x_i + e_i), \quad (27)$$

kde x_i jsou vysvětlující proměnné (též regresory), β jsou regresní koeficienty, které je třeba odhadnout a e_i jsou nekorelované náhodné chyby [15]. Aby mohl být lineární regresní model aplikován musí být splněny následující předpoklady:

- 1) Náhodná veličina má nulovou střední hodnotu

$$E(e_i) = 0, \text{ pro } \forall i \in I \quad (28)$$

- 2) Náhodná veličina má konstantní kladný rozptyl

$$D(e_i) = \sigma^2 > 0, \text{ pro } \forall i \in I \quad (29)$$

- 3) Náhodné veličiny Y_i pochází z normálního rozdělení pro $i = 1, \dots, n$ [15] a [30].

Pokud datový soubor náhodných veličin splňuje výše uvedené předpoklady, mohou být odhadnuty koeficienty regresní rovnice pomocí metody

nejmenších čtverců. Při této metodě je třeba v prvním kroku sestavit matici ze získaných dat

$$\mathbf{H}_1 = \begin{pmatrix} \sum_{i=1}^n x_{1i}x_{1i} & \cdots & \sum_{i=1}^n x_{1i}x_{mi} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{mi}x_{1i} & \cdots & \sum_{i=1}^n x_{mi}x_{mi} \end{pmatrix}, \quad (30)$$

kde n je počet členů souboru a m je hodnost matice.

Aby bylo možné sestavit rovnici přímky, musejí být určeny bodové a intervalové odhad regresních koeficientů β_i , které odpovídají matici \mathbf{b} . Hodnoty matice \mathbf{b} se vypočítají jako soustava lineárních rovnic

$$\mathbf{H}_1 \mathbf{b} = \mathbf{g}, \quad (31)$$

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, \quad (32)$$

$$\mathbf{g} = \begin{pmatrix} \sum_{i=1}^n x_{1i}y_i \\ \vdots \\ \sum_{i=1}^n x_{mi}y_i \end{pmatrix}. \quad (33)$$

Nyní je možné určit rovnici přímky a také minimální reziduální hodnotu součtu čtverců

$$S_{min}^* = \sum_{i=1}^n y_i^2 - \sum_{j=1}^m b_j g_j \quad (34)$$

pomocí kterého se vypočítá bodový odhad rozptylu σ^2

$$s^2 = \frac{S_{min}^*}{n - m} \quad (35)$$

a směrodatná odchylka dle vzorce (9).

Kvalitu regresní funkce vyjadřuje koeficient vícenásobné korelace

$$r = \sqrt{1 - \frac{S_{min}^*}{\sum y_i^2 - n(\bar{y})^2}}, \quad (36)$$

kde r nabývá hodnot z intervalu $\langle 0, 1 \rangle$ [15].

3 DATOVÁ SADA

Datové sady, na základě kterých jsou prováděny výpočty, byly získány roztríděním SKO z černých popelnic. Třídění odpadu probíhalo v rámci projektu TIRSMZP719.

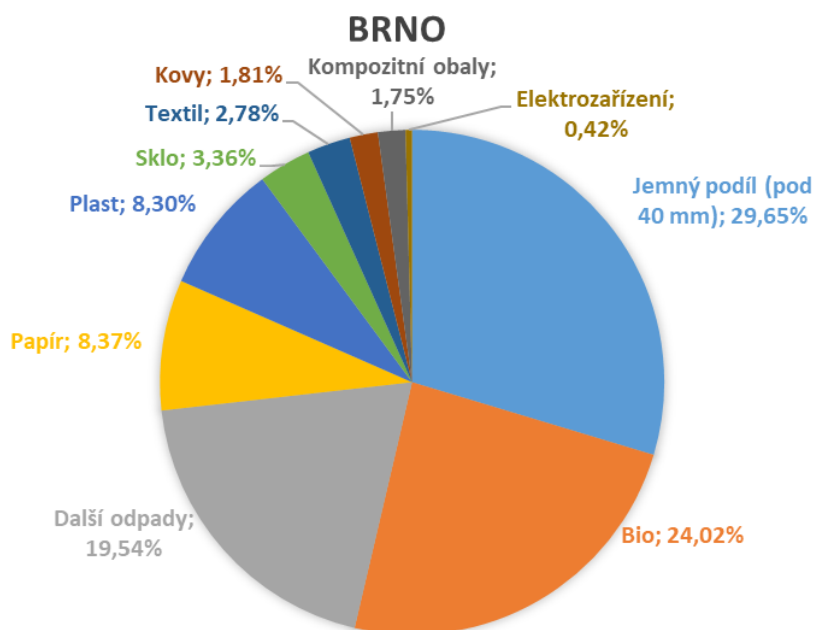
Rozbory byly prováděny pro odpady z Brna (statutární město) obr. 8, Valašských Klobouků (spádové město obce s rozšířenou působností) obr. 9 a Louky (obec) obr. 10. Data jsou dále rozdělena podle typů zástavby. V Brně je to typ venkovské, sídlištní a vilové zástavby, ve Valašských Kloboucích typ sídlištní a vilové a v Louce je to pouze typ venkovský. Základní údaje o získaných datech pro vybrané odpadové frakce jsou zapsány v tab. 2.

Tab. 2 Hodnoty datových sad pro jednotlivé frakce

Lokalita Zástavba	Váha [kg]	Počet vzorků	Naměřené a vypočítané hodnoty					
			Papír		Plast		Bio	
			Průměr	Odchylka	Průměr	Odchylka	Průměr	Odchylka
			[%]	[%]	[%]	[%]	[%]	[%]
Brno Venkovská	619	3	6,76	1,03	7,25	0,32	27,54	6,21
Brno Sídlištní	2497	13	10,14	1,01	8,90	0,76	23,37	2,16
Brno Vilová	1355	7	8,21	1,67	8,75	0,80	21,13	4,60
Brno Celkem	4471	23	8,37	0,74	8,30	0,38	24,02	2,67
Valašské Klobouky Sídlištní	805	13	4,96	0,77	4,38	0,57	33,00	3,78
Valašské Klobouky Vilová	1980	45	3,76	0,65	3,93	0,31	28,05	3,97
Valašské Klobouky Celkem	2785	58	4,36	0,50	4,16	0,33	30,53	2,74
Louka Venkovská	481	43	5,28	1,24	5,03	0,86	42,38	5,14

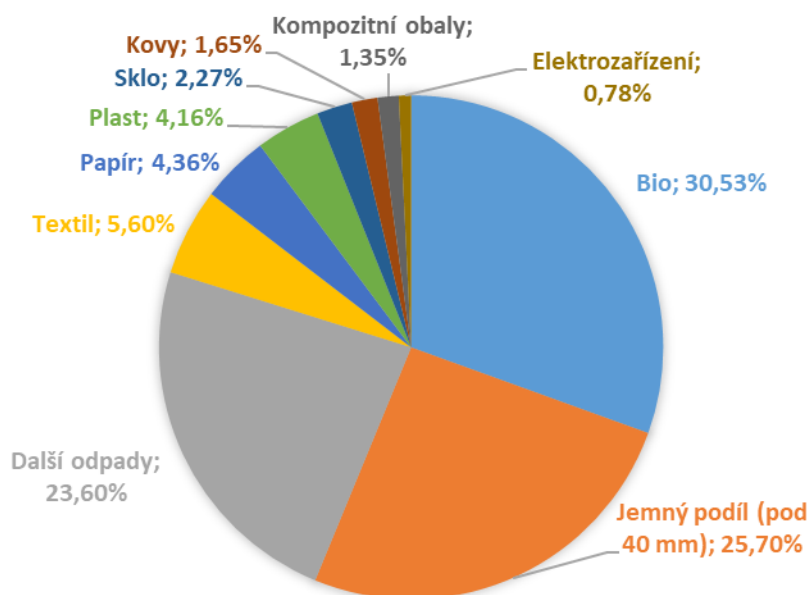
Lokality pro porovnávání dat byly vybrány na základě stratifikace obcí. V rámci ní byly vybráni vhodní reprezentanti pro jednotlivá strata. Tímto rozdělením se zabývala odborná studie *Vývoj statistické metodiky pro místní šetření – směsného komunálního odpadu na úrovni producentů odpad* v rámci projektu SPETEP. Tato studie rozdělila ČR do strat, kde každé stratum obsahuje obce s podobnými charakteristikami. Zatím není dostupná potřebná datová sada (pouze tři pozorování), a proto je možné provádět pouze dílčí testování, které však mohou naznačit, zda byl výběr lokalit na základě expertně vybraných faktorů správný. Vybrané proměnné byly např. počet obyvatel, produkce SKO [kg/obyv.],

hustota zalidnění v zástavbě [obyv./ha], průměr počtu obyvatel na adresu, produkce separovaného plastu, papíru, skla a bioodpadu aj. Každá lokalita reprezentuje určitý druh obce. V obcích jsou dále testovány rozdílnosti výsledky složení odpadu vzhledem k různým typům zástavby. Testování je prováděno z malé, avšak pro tyto účely postačující datové sady. Jak již bylo zmíněno, datový základ je omezen kvůli nepříznivé pandemické situaci v ČR v době vypracování této práce. Přístup je však obecně použitelný opakovaně pro plávanou zvětšující se datovou sadu v budoucnu.



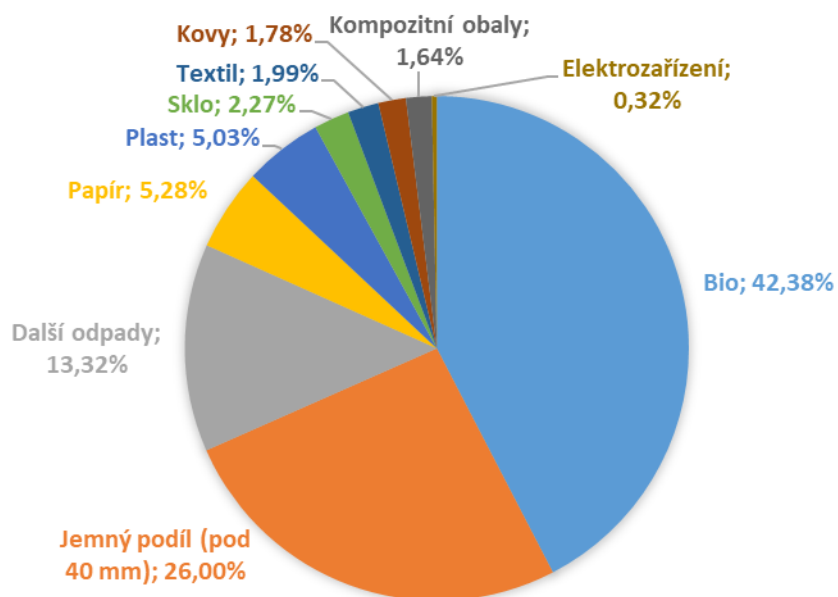
Obr. 8 Průměrné složení odpadu Brna

VALAŠSKÉ KLOBOUKY



Obr. 9 Průměrné složení odpadu Valašských Klobouků

LOUKA



Obr. 10 Průměrné složení odpadu Louky

4 POSTUP ZPRACOVÁNÍ DAT A PŘEDPOKLADY

Vyhodnocování a porovnávání datových sad probíhá v několika krocích. Schéma průběhu vyhodnocování testů je znázorněno na obr. 11. Postup použití jednotlivých testů je ovlivněn testy jimi předcházejícími. U dat jsou porovnávány obce jako celky a jednotlivé typy zástaveb. Pro obě tato porovnávání jsou použity stejné statistické testy.

V prvním kroku postupu je proveden Kolmogorovův-Smirnovův test (viz kapitola 2.6 *Testování hypotéz*). Testem je ověřena normalita jednotlivých datových sad. Tento test se provádí pro všechny datové sady a podle jeho výsledků jsou provedeny následující kroky postupu.

1. Data nepochází z normálního rozdělení

Jestliže při porovnávání dvou datových sad nepochází alespoň jedna sada z normálního rozdělení, poté je třeba na porovnávání sad použít neparametrický test. Dále je tedy aplikován dvouvýběrový Wilcoxonův test (viz kapitola 2.6 *Testování hypotéz*). Tento test ukazuje, zda datové sady mají podobné vlastnosti (střední hodnoty) a na tento test již v této práci žádný další nenavazuje.

2. Data pochází z normálního rozdělení

Pokud obě datové sady mají normální rozdělení, jsou dále porovnávány parametrickými testy. Tyto testy mají oproti neparametrickým testům větší sílu, tj. menší chybu 2. druhu. Tato práce se především zabývá středními hodnotami a rozptyly porovnávaných datových sad. Dalším krokem je tedy zjistit, zda mají datové sady podobné rozptyly. Tuto vlastnost je možné testovat pomocí Fisherova testu.

Dalším testem je Studentův test, kterým lze posoudit shodnost středních hodnot datových sad. Studentův test má více variant, ale v této práci je použit pouze Studentův test pro dva výběry. Při tomto testu jsou použity jeho dvě varianty.

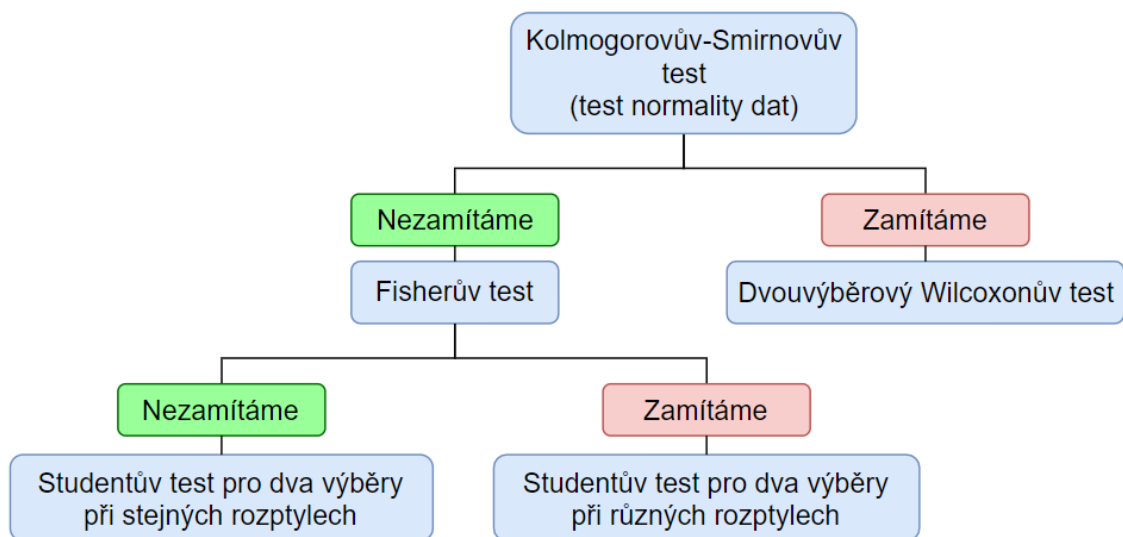
2. 1 Hypotézu o shodnosti rozptylů nebyl zamítnut

Fisherův test nebyl zamítnut, a proto je v následujícím kroku použit Studentův test při stejných rozptylech.

2. 2 Hypotézu o shodnosti rozptylů byl zamítnut

Fisherův test byl zamítnut a dále tedy je použit Studentův test při různých rozptylech (obě zmíněné varianty jsou blíže popsány v kap. 2.6 *Testování hypotéz*).

Použité Studentovy testy ukazují, zda mají nebo nemají datové sady statisticky stejné střední hodnoty, což je cílem prováděných analýz.



Obr. 11 Schéma znázornění postupu vyhodnocování hypotéz

5 VYHODNOCOVÁNÍ HYPOTÉZ

Pro vyhodnocování a určení podobných vlastností datových sad (shodnost středních hodnot) byly použity testy, které jsou uvedeny v kapitole 2.6 *Testování hypotéz*. Výpočty byly provedeny v několika krocích viz. obr. 11.

Jako první krok k ověřování hypotéz bylo třeba určit, zda data pocházejí z normálního rozdělení. Normalita dat se testuje pomocí Kolmogorovova-Smirnovova testu popsaného v kapitole 2.6 *Testování hypotéz*. U tohoto testu je výstupem hodnota D_n , která je porovnávána s kritickou hodnotou $D_n(\alpha)$. Pokud platí, že hodnota $D_n < D_n(\alpha)$, pak na hladině významnosti $\alpha = 0,05$ není hypotéza zamítnuta a s datovou sadou je dále pracováno jako by pocházela z normálního rozdělení. Pro tyto statistické analýzy byly vybrány tři základní frakce SKO, konkrétně papír, plast, bio odpad.

Tab. 3 Výsledky pro normalitu dat

Lokalita	Zástavba	Odpovídá normálnímu rozdělení								
		Papír			Plast			Bio		
		D_n	$D_n(\alpha)$		D_n	$D_n(\alpha)$		D_n	$D_n(\alpha)$	
Brno	Venkovská	0,274	0,708	N	0,191	0,708	N	0,250	0,708	N
Brno	Sídlště	0,197	0,361	N	0,121	0,361	N	0,187	0,361	N
Brno	Vilová	0,197	0,483	N	0,139	0,483	N	0,151	0,483	N
Brno	Celkem	0,178	0,275	N	0,083	0,275	N	0,061	0,275	N
Valašské Klobouky	Sídlště	0,138	0,361	N	0,167	0,361	N	0,183	0,361	N
Valašské Klobouky	Vilová	0,113	0,198	N	0,139	0,198	N	0,079	0,198	N
Valašské Klobouky	Celkem	0,091	0,175	N	0,114	0,175	N	0,095	0,175	N
Louka	Venkovská	0,189	0,203	N	0,128	0,203	N	0,085	0,203	N

Pozn.: Z – zamítáme hypotézu; N – nezamítáme hypotézu

Výsledky Kolmogorovova-Smirnovova testu ukazují, že datové sady pro jednotlivé zástavby i obce jako celky pocházejí z normálního rozdělení. Jelikož všechny hypotézy tohoto testu byly nezamítnuty, nebude v následujících výpočtech použit dvouvýběrový wilcoxonův test.

Pro zástavby a obce je tedy v následujícím kroku aplikován Fisherův test pro určení shodnosti rozptylů datových sad. Fisherův test je použit k porovnávání datových sad pro všechny typy zástaveb mezi sebou, tedy ke srovnání zástaveb v jednotlivých obcích a stejných typů zástaveb v různých obcích. Tímto testem jsou také porovnány obce jako celky. V tab. 4 jsou uvedeny všechny výsledky Fisherova testu. Aby hypotéza nebyla zamítnuta musí hodnota t náležet intervalu přijetí \bar{W}_α .

Tab. 4 Výsledky Fisherova testu

Lokalita	Druh odpadu	Porovnávaná hodnota t	Interval pro přijetí hypotézy \bar{W}_α	Přijetí hypotézy
Brno: venkovská – sídlištní	Papír	1,439	<1; 5,096>	Nezamítáme
	Plast	4,052	<1; 39,415>	Nezamítáme
	Bio	11,469	<1; 5,096>	Zamítáme
Brno: venkovská – vilová	Papír	2,047	<1; 39,331>	Nezamítáme
	Plast	4,822	<1; 39,331>	Nezamítáme
	Bio	2,341	<1; 7,260>	Nezamítáme
Brno: sídlištní – vilová	Papír	2,946	<1; 3,728>	Nezamítáme
	Plast	1,190	<1; 3,728>	Nezamítáme
	Bio	4,900	<1; 3,728>	Zamítáme
Valašské Klobouky: sídlištní – vilová	Papír	1,462	<1; 2,256>	Nezamítáme
	Plast	3,593	<1; 2,256>	Zamítáme
	Bio	1,039	<1; 2,894>	Nezamítáme
Brno – Valašské Klobouky: Sídlištní	Papír	1,726	<1; 3,277>	Nezamítáme
	Plast	1,755	<1; 3,277>	Nezamítáme
	Bio	3,069	<1; 3,277>	Nezamítáme
Brno – Valašské Klobouky vilová	Papír	7,431	<1; 2,713>	Zamítáme
	Plast	7,503	<1; 2,713>	Zamítáme
	Bio	1,536	<1; 2,713>	Nezamítáme
Brno – Louka: venkovská	Papír	1,001	<1; 4,034>	Nezamítáme
	Plast	4,951	<1; 39,474>	Nezamítáme
	Bio	2,136	<1; 4,034>	Nezamítáme
Brno – Valašské Klobouky	Papír	2,182	<1; 1,927>	Zamítáme
	Plast	1,413	<1; 1,927>	Nezamítáme
	Bio	1,021	<1; 2,155>	Nezamítáme
Brno – Louka	Papír	2,808	<1; 2,204>	Zamítáme
	Plast	4,991	<1; 2,204>	Zamítáme
	Bio	3,620	<1; 2,204>	Zamítáme
Valašské Klobouky – Louka	Papír	6,127	<1; 1,751>	Zamítáme
	Plast	7,052	<1; 1,751>	Zamítáme
	Bio	3,546	<1; 1,751>	Zamítáme

Fisherův test byl především zamítnut u porovnávání celých obcí. Ukázalo se, že datové sady mají podobné rozptyly spíše při porovnávání zástaveb než při porovnávání celých obcí.

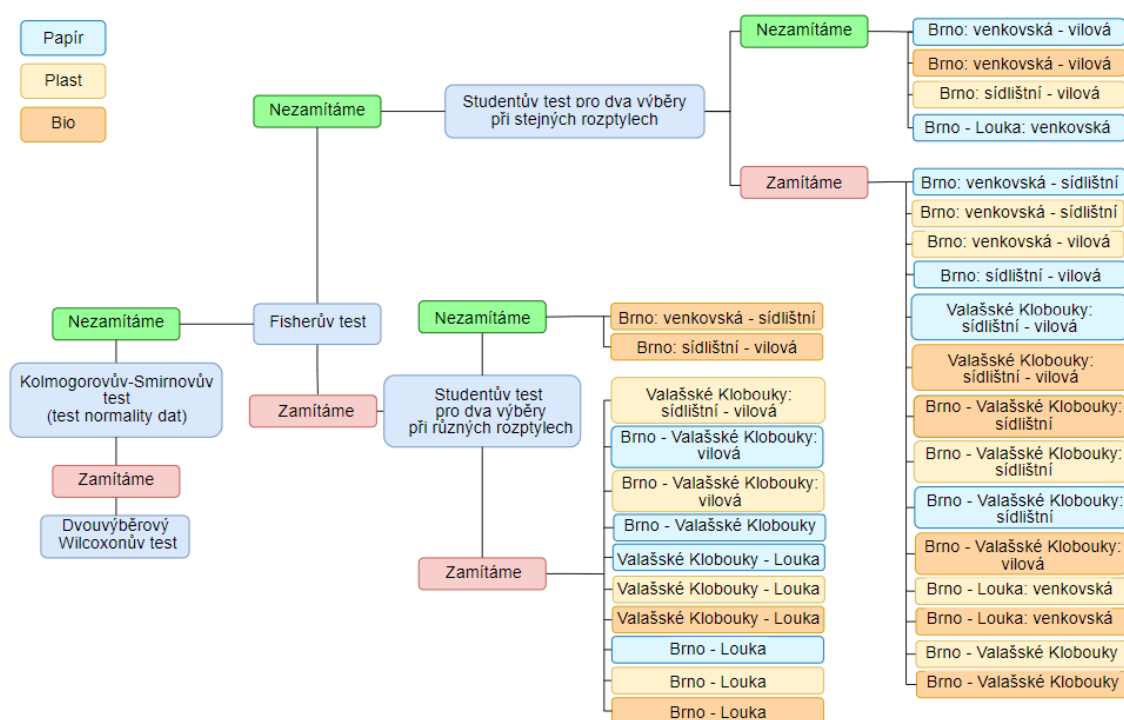
Pro lokality, u kterých nebyl test zamítnut, je dále použit Studentův test pro dva výběry při stejných rozptylech. U lokalit, kde byl Fisherův test zamítnut, je dále použit Studentův test pro dva výběry při různých rozptylech. U některých

porovnávaných dat vyšel velký interval pro přijetí hypotézy, což mohlo ovlivnit její potvrzení či vyvrácení. Velikost intervalu je ovlivněna především malým množstvím dostupných pozorování. Výsledky pro obě varianty t-testu jsou uvedeny v tab. 5.

Tab. 5 Výsledky Studentova testu pro stejné/různé rozptyly

Lokalita	Druh odpadu	Porovnávaná hodnota t	Interval pro přijetí hypotézy \bar{W}_α	Přijetí hypotézy
Brno: venkovská – sídlištní	Papír	4,875	<-2,145; 2,145>	Zamítáme
	Plast	3,488	<-2,145; 2,145>	Zamítáme
	Bio	0,941	<-4,261; 4,261>	Nezamítáme
Brno: venkovská – vilová	Papír	1,250	<-2,306; 2,306>	Nezamítáme
	Plast	2,808	<-2,306; 2,306>	Zamítáme
	Bio	1,618	<-2,306; 2,306>	Nezamítáme
Brno: sídlištní – vilová	Papír	3,051	<-2,101; 2,101>	Zamítáme
	Plast	0,401	<-2,101; 2,101>	Nezamítáme
	Bio	1,133	<-2,420; 2,420>	Nezamítáme
Valašské Klobouky: sídlištní – vilová	Papír	5,480	<-2,004; 2,004>	Zamítáme
	Plast	2,618	<-2,167; 2,167>	Zamítáme
	Bio	3,935	<-2,004; 2,004>	Zamítáme
Brno – Valašské Klobouky: sídlištní	Papír	14,149	<-2,064; 2,064>	Zamítáme
	Plast	16,461	<-2,064; 2,064>	Zamítáme
	Bio	7,666	<-2,064; 2,064>	Zamítáme
Brno – Valašské Klobouky: vilová	Papír	6,460	<-2,438; 2,438>	Zamítáme
	Plast	14,634	<-2,438; 2,438>	Zamítáme
	Bio	4,118	<-2,009; 2,099>	Zamítáme
Brno – Louka: venkovská	Papír	1,970	<-2,015; 2,015>	Nezamítáme
	Plast	4,874	<-2,015; 2,015>	Zamítáme
	Bio	4,657	<-2,064; 2,064>	Zamítáme
Brno – Valašské Klobouky	Papír	12,474	<-2,041; 2,041>	Zamítáme
	Plast	48,426	<-1,990; 1,990>	Zamítáme
	Bio	9,593	<-1,990; 1,990>	Zamítáme
Brno – Louka	Papír	12,474	<-2,041; 2,041>	Zamítáme
	Plast	20,964	<-2,033; 2,033>	Zamítáme
	Bio	18,798	<-2,037; 2,037>	Zamítáme
Valašské Klobouky – Louka	Papír	4,506	<-2,017; 2,017>	Zamítáme
	Plast	12,474	<-2,017; 2,017>	Zamítáme
	Bio	13,589	<-2,016; 2,016>	Zamítáme

Výsledky Studentova testu ukázaly, že stejné střední hodnoty mají spíše porovnávané zástavby v rámci jedné obce. Pro porovnávání stejných zástaveb v různých obcích byly téměř všechny testy zamítnuty. Při porovnávání zástaveb v rámci obce se ukázalo, že možná záleží na velikosti obce. Ve Valašských Kloboukách byly zamítnuty oba testy, v Brně byla zamítnuta méně než polovina testů. Mohlo by to naznačit, že ve větších městech bude větší pravděpodobnost shody středních hodnot porovnávaných zástaveb. Pro jasné prokázání této úvahy by bylo třeba provést další testování na rozšířené datové sadě. Porovnávání obcí jako celku ukázalo, že sledované obce mají odlišné střední hodnoty. Celkové výsledky testů o shodě středních hodnot pro papír, plast a bio odpad jsou sumarizovány na obr. 12.



Obr. 12 Schéma výsledků provedených testů

Z obr. 12 je patrné, že nejvíce se podobají typy zástaveb: venkovská – vilová (papír a bio odpad) a sídlištní – vilová (plast a bio odpad) v Brně. Z výsledků je možné dále vidět, že při testování nebudou nějak výrazné rozdíly mezi jednotlivými frakcemi SKO, tedy mezi papírem, plastem a bio odpadem. Tyto tři zástupci SKO jsou ve výsledku rovnoměrně rozloženy a bude spíše záležet na porovnávaných oblastech.

Testy pro celé obce byly z velké části zamítnuty, což by mohlo naznačit, že výběr reprezentantů se zdá být na základě této dílčí analýzy správný. Již z těchto testů lze konstatovat, že reprezentativní obce na základě stratifikace byly vybrány správně a všechny přispívají k popisu variability složení SKO v rámci ČR. Naopak v případě stratifikace v obcích na základě typu zástavby

se správnost rozdělení vykazuje pouze v některých případech a je možné, že s přibývajícími daty bude vhodné rozdělování obcí na podoblasti přehodnotit.

6 POUŽITÍ KORELAČNÍ A REGRESNÍ ANALÝZY

S ohledem na omezené množství dat z rozborů odpadů vlivem pandemie Covid-19 není možné provést na existující datové sadě korelační a regresní analýzu. Vzhledem k plánovanému dalšímu využití prezentovaného přístupu, na v budoucnu rozšířené datové sadě, je v této kapitole popsán postup využití korelační a regresní analýzy alespoň teoreticky. Jsou zde uvedené podstatné kroky analýz včetně doporučení autorky této práce.

Pomocí korelační analýzy je možné vytipovat vlivné faktory na složení SKO. Tento krok je možný až v případě dostupné datové sady z rozborů odpadu na konci projektu TIRSMZP719 (řádově 10 pozorování). Pro stratifikaci v prvním kroku bylo nutné vybrat vlivné faktory na základě expertního názoru. V rámci projektu SPETEP byly expertně stanoveny následující faktory:

- Ekonomická data: počet obyvatel dojíždějících za prací do obce, počet dojíždějících za prací mimo obec.
- Sociální data: počet obyvatel, hustota zalidnění v zastavěné části obce, hustota zalidnění na celém území obce, průměrný počet obyvatel na číslo popisné, podíl obydlených bytů v rodinných domech z celkového počtu, celkový počet ubytovacích zařízení, počet neobydlených bytů určených k rekreaci.
- Odpadové data: produkce separovaného plastu, produkce separovaného papíru, produkce separovaného skla, produkce separovaného bio odpadu, produkce SKO.
- Jiná data: přípojka k ústřednímu topení.

Autorka práce dále navrhuje rozšířit datovou sadu pro tuto analýzu. Díky tomu by mohlo být dosaženo lepší stratifikace (popisu variability) a následného modelu založeného na regresní analýze. Limitem jsou dostupná data na úrovni obce. Důležitý zdroj dat poskytuje v této oblasti Český statistický úřad. Konkrétně je možné zmínit datovou sadu s názvem *Územně analytické podklady* [31]. Ale lze najít i další cenné data pro sledovaný parametr – průměrné složení SKO. Níže jsou uvedené některé vybrané faktory:

- Ekonomická data: míra nezaměstnanosti, podíl lidí pracujících v zemědělství aj.
- Sociální data: průměrný věk, charakter vzdělání, počet studentů vysokých škol v obci aj.
- Odpadová data: celková produkce KO, produkce separovaného textilu, míra separace vybraných odpadů aj.
- Jiná data: způsob vytápění aj.

V prvním kroku je třeba vytvořit korelační matici na základě vzorců (24) nebo (25). Zde je třeba zvážit lineární vs. nelineární vazby mezi závislou

a nezávislou proměnnou. V případě použití Pearsnova korelačního koeficientu je třeba otestovat normalitu dat. Před samotnou analýzou je vhodné odstranit z dat odlehlé hodnoty. Vzhledem k pravděpodobnému charakteru dat (nesplnění normality, hodně odlehlých hodnot) je vhodné využít robustního přístupu. Autorka navrhuje použití následujícího vzorce publikovaného v [32].

$$Q = \frac{x_i - \text{median}(x_j)}{1,483 \text{median}|x_i - \text{median}(x_j)|} \quad (37)$$

kde x_i jsou posuzovaná data a parametr Q je testovací kritérium, které je dle autorů [32] vhodné nastavit na prahovou hodnotu tři (pro hodnotu Q větší než tři jsou identifikovaná data jako odlehlá pozorování). Dále je vhodné vynechat z analýzy extrémní hodnoty. Tu představuje např. Hl. m. Praha, která se svým charakterem významně liší od zbytku obcí. Dále to jsou malé vesničky, které jsou rozsáhlé do počtu, ale nepředstavují velký podíl na celkové produkci SKO. Autorka doporučuje vynechat spodních 2,5 % produkovaného SKO prostřednictvím agregace nejmenších obcí z pohledu produkce odpadu.

V druhém kroku, když jsou identifikovány významné faktory je možné vytvořit lineární regresní model viz *2.7 Korelační a regresní analýza*, který se doporučuje v momentě, kdy není znám nějaký fyzikální nebo jiný důvod pro složitější vztah (nelineární) regresní analýzy. Jestliže se podaří nalézt statisticky významný model, je díky němu možné odhadnou složení SKO pro všechny obce, kde nebyly provedeny rozborů složení. Zároveň je ale nutné disponovat daty o faktorech užitých u stratifikace a následné regresní analýzy ve všech obcích, pro které je cílem odhad průměrného složení, na základě modelu, udělat.

7 SHRNUTÍ A ZÁVĚRY

Produkce odpadu neodmyslitelně patří k životu člověka a je spojena téměř s veškerou jeho činností. Většina KO má velký potenciál pro následné využití, jen je třeba jej dobře připravit k následnému zpracování. Pro tuto separaci jsou v obcích umístěny barevné kontejnery na recyklovatelný materiál. Ne všechen odpad z domácností je však roztříděn a mnoho roztříditelného materiálu končí v černých směsných komunálních kontejnerech. Aby bylo možné hledat cestu k efektivnějšímu nakládání s KO, je třeba mít představu o složení zbytkové frakce, tj. SKO.

V rámci projektu TIRSMZP719 proběhlo zkoumání složení SKO v několika lokalitách. Důležitým návazným krokem je statistické testování získaných dat. Zpracování dat se v této práci zaměřilo na podíl papíru, plastu a bio odpadu v SKO. Jak je znázorněno v grafech na obr. 8, obr. 9 a obr. 10, plast s papírem tvoří kolem 4–8 % SKO a bio odpad tvoří podstatně větší část, a to mezi 20–40 % SKO. Tyto procentuální hodnoty by se daly minimalizovat, kdyby byl odpad umístěn do správných kontejnerů pro separovaný odpad. Prvním krokem ke zlepšení aktuální situace je znalost průměrného složení v černých kontejnerech. A to nejen pro stát jako celek, ale i pro detailnější územní členění.

Pro lepší vhled do dostupné datové sady byla provedena dílčí statistická analýza. Rozsah analýzy byl dán pozastavením rozborů odpadu (sběru dat) vlivem pandemie COVID-19. Jádrem této bakalářské práce bylo testování hypotéz za účelem zjištění správnosti stratifikace obcí a sběrných nádob.

Výsledky statistických testů ukazují na to, že výběr reprezentantů na úrovni obcí byl správný. Téměř všechna testování o rozdílnosti středních hodnot byla zamítnuta, což ukazuje na to, že na složení SKO bude mít vliv rozdílný charakter obcí. Testování hypotéz na úrovni zástaveb v jednotlivých obcích bylo více než z poloviny nezamítnuto. Tyto výsledky naznačují, že v rámci obce nebudou takové rozdíly obsahu v SKO. Bude tedy třeba provést další testování, aby tyto hypotézy byly potvrzeny na rozsáhlejší datové sadě. Podstatným výsledkem by byl fakt, že v rámci obce nezáleží na jeho dílčích částech (dané např. typem zástavby), ale je možné ji brát jako celek. Tento aspekt by výrazně zlevnil rozbor odpadu z důvodu nižšího počtu potřebných pozorování. Posledními porovnávanými lokalitami byly stejné typy zástaveb v různých obcích. Testy ukázaly, že jejich střední hodnota se liší. Rozdílné střední hodnoty opět ukazují na fakt, že je třeba brát obce jako celek a nekombinovat stejné charakteristiky zástavby mezi více obcemi. Tento aspekt je podstatný především pro následné konstrukce regresních a jiných modelů.

V případě, že se podaří rozšířit aktuální datovou sadu, bude možné provést korelační a následně regresní analýzu. To umožní ověřit správnost stratifikace a případně navrhnout vhodné úpravy, které zajistí popis většího množství variability ve složení SKO v jednotlivých obcích ČR. To je zásadní

aspekt pro dosažení kvalitního odhadu průměrného složení jak pro agregovaná data (stát kraj aj.), tak pro jednotlivé obce. Dalším přínosem regresní analýzy je odhad složení SKO i pro obce, kde rozbor odpadu nebyl proveden. Nutností je disponovat v těchto obcích stejnými vysvětlujícími daty, které byly využity v regresní funkci.

SEZNAM POUŽITÉ LITERATURY

- [1] Česká republika. Zákon č. 185/2001 Sb., ze dne 15. května 2001, o odpadech a o změně některých dalších zákonů. In: *Sbírka zákonů České republiky*. 2001, částka 71, s. 4074-4113. ISSN 1211-1244. Dostupné také z: <https://www.zakonyprolidi.cz/cs/2001-185>
- [2] FERDAN, Tomáš, Martin PAVLAS, Vlastimír NEVRILÝ, Radovan ŠOMPLÁK a Petr STEHLÍK. Greenhouse gas emissions from thermal treatment of non-recyclable municipal waste. *Frontiers of Chemical Science and Engineering* [online]. 2018, 12(4), 815-831 [cit. 2021-2-21]. ISSN 2095-0179. DOI: 10.1007/s11705-018-1761-4 Dostupné z: <https://doi.org/10.1007/s11705-018-1761-4>
- [3] Česká republika. Zákon č. 541/2020 Sb., ze dne 1. prosince 2020, o odpadech. In: *Sbírka zákonů České republiky*. 2020, částka 222, s. 6082–6185. ISSN 1211-1244. Dostupné také z: <https://www.zakonyprolidi.cz/cs/2020-541#cast3>
- [4] Směrnice Evropského parlamentu a Rady (EU) 2018/851 ze dne 30. května 2018, kterou se mění směrnice 2008/98/ES o odpadech. In: *Úřední věstník*, L 150, 14. 6. 2018, s. 109–140. Dostupné také z: <https://eur-lex.europa.eu/legal-content/CS/TXT/?uri=CELEX%3A32018L0851>
- [5] Směrnice Evropského parlamentu a Rady (EU) 2018/852 ze dne 30. května 2018, kterou se mění směrnice 94/62/ES o obalech a obalových odpadech. In: *Úřední věstník*, L 150, 14. 6. 2018, s. 141–154. Dostupné také z: <https://eur-lex.europa.eu/legal-content/CS/TXT/?uri=CELEX%3A32018L0852>
- [6] Směrnice Evropského parlamentu a Rady (EU) 2018/850 ze dne 30. května 2018, kterou se mění směrnice 1999/31/ES o skládkách odpadů. In: *Úřední věstník*, L 150, 14. 6. 2018, s. 100–108. Dostupné také z: <https://eur-lex.europa.eu/legal-content/CS/TXT/?uri=celex:32018L0850>
- [7] Odpadové a oběhové hospodářství. *Informační systém odpadového hospodářství ISOH* [online]. CENIA [cit. 2021-2-21]. Dostupné z: <https://www.cenia.cz/odpadove-a-obehove-hospodarstvi/isoh/>
- [8] PLUSKAL, Jaroslav, Radovan ŠOMPLÁK, Vlastimír NEVRILÝ, Veronika SMEJKALOVÁ a Martin PAVLAS. Strategic decisions leading to sustainable waste management: Separation, sorting and recycling possibilities. *Journal of Cleaner Production* [online]. 2021, 278 [cit. 2021-2-21]. ISSN 09596526. DOI: 10.1016/j.jclepro.2020.123359 Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S0959652620334041>

- [9] *Methodology for the Analysis of Solid Waste* [online]. (SWA-Tool), 2004 [cit. 2021-2-21]. Dostupné z: <https://www.wien.gv.at/meu/fdb/pdf/swa-tool-759-ma48.pdf>
- [10] *Skladba směsného komunálního odpadu z domácností ČR* [online]. EKO-KOM, 2019 [cit. 2021-2-21]. Dostupné z: <https://www.ekokom.cz/news/715/212/Skladba-smesneho-komunalniho-odpadu-z-domacnosti-cR>
- [11] *Skladba domovního odpadu SKO v České Republice v roce 2016. Zpravodaj: Elektronický zpravodaj pro města a obce systému EKO-KOM* [online]. Praha 4, 2017, (18), 3 [cit. 2021-2-21]. Dostupné z: https://www.ekokom.cz/uploads/news/id627/ZPR@VODAJ_18_2017_fi nal.pdf
- [12] TIRSMZP719 Prognózování produkce odpadů a stanovení složení komunálního odpadu, TAČR, doba řešení: 1.1.2019–31.12.2021.
- [13] SPETEP Strategické partnerství pro environmentální technologie a produkci energie, poskytovatel: Ministerstvo školství, mládeže a tělovýchovy České republiky, CZ.02.1.01/0.0/0.0/16_026/0008413, doba řešení: 01.07.2018 — 30.06.2022
- [14] ANDĚL, Jiří. *Matematická statistika*. Praha: SNTL – Nakladatelství technické literatury, 1987.
- [15] KARPÍŠEK, Zdeněk. *Matematika IV: Statistika a pravděpodobnost*. 3., dopl. vyd. Brno: Akademické nakladatelství CERM, 2007. ISBN 978-80-214-3380-9.
- [16] *Matematika online: Matematika IV* [online]. [cit. 2021-1-29]. Dostupné z: <https://mathonline.fme.vutbr.cz/Matematika-IV/sc-108-sr-1-a-120/default.aspx>
- [17] BEDNÁŘOVÁ, Iveta. Náhodná veličina. *Biostatistika: Bezpečnost a kvalita potravin* [online]. Brno: Veterinární a farmaceutická univerzita [cit. 2021-2-15]. Dostupné z: <https://cit.vfu.cz/statpotr/POTR/Teorie/Predn1/nahvelic.htm>
- [18] BUDÍKOVÁ, Marie. Centrální limitní věta. *Statistika a pravděpodobnost* [online]. Brno: Přírodovědecká fakulta Masarykovy univerzity, 2016 [cit. 2021-2-5]. Dostupné z: <https://is.muni.cz/do/rect/el/estud/prif/ps15/statistika/web/pages/centralni-limitni-veta.html>

- [19] HOLČÍK, Jiří, Martin KOMENDA a kol. Normální rozdělení pravděpodobnosti. *Matematická biologie: e-learningová učebnice* [online]. Brno: Masarykova univerzita, 2015 [cit. 2021-2-5]. Dostupné z: <https://portal.matematickabiologie.cz/index.php?pg=aplikovana-analyza-klinickyh-a-biologickyh-dat--biostatistika-pro-matematickou-biologii--nahodna-velicina-rozdeleni-pravdepodobnosti-a-realna-data--normalni-rozdeleni-pravdepodobnosti>
- [20] DORDA, Michal. *Odhady parametrů základního souboru* [online]. [cit. 2021-02-25]. Dostupné z: http://homel.vsb.cz/~dor028/Aplikace_6.pdf
- [21] NEUBAUER, Jiří. *Základy teorie odhadu parametrů - bodový odhad* [online]. [cit. 2021-2-25]. Dostupné z: https://k101.unob.cz/~neubauer/pdf/bodovy_odhad.pdf
- [22] Bodové a intervalové odhady parametrů rozdělení pravděpodobnosti. *Maths and Stats Support Centre* [online]. Katedra aplikované matematiky a informatiky Ekonomicko-správní fakulty Masarykovy univerzity [cit. 2021-5-3]. Dostupné z: https://mathstat.econ.muni.cz/media/19023/bodove_odhady.pdf
- [23] HOLČÍK, Jiří, Martin KOMENDA a kol. Bodové a intervalové odhady parametrů normálního rozdělení. *Matematická biologie: e-learningová učebnice* [online]. Brno: Masarykova univerzita, 2015 [cit. 2021-3-5]. Dostupné z: <https://portal.matematickabiologie.cz/index.php?pg=analiza-a-hodnoceni-biologickyh-dat--statisticke-modelovani--zakladni-pojmy-matematicke-statistiky--bodove-a-intervalove-odhady-parametru-normalniho-rozdeleni>
- [24] PAVLÍK, Jiří. *Aplikovaná statistika* [online]. Praha: Vysoká škola chemicko-technologická, 2005 [cit. 2021-3-5]. ISBN 80-7080-569-2. Dostupné z: http://147.33.74.135/knihy/uid_isbn-80-7080-569-2/pages-img/001.html
- [25] HOLČÍK, Jiří, Martin KOMENDA a kol. Kolmogorovův - Smirnovův test. *Matematická biologie: e-learningová učebnice* [online]. Brno: Masarykova univerzita, 2015 [cit. 2021-3-5]. Dostupné z: <https://portal.matematickabiologie.cz/index.php?pg=analiza-a-hodnoceni-biologickyh-dat--statisticke-modelovani--overovani-predpokladu-v-klasickem-modelu-linearni-regrese--overovani-normality-dat--kolmogorovuv-smirnovuv-test>

- [26] BEDÁŇOVÁ, Iveta. Parametrické testy – Studentův t-test. *Biostatistika: Bezpečnost a kvalita potravin* [online]. Brno: Veterinární a farmaceutická univerzita [cit. 2021-3-10]. Dostupné z: <https://cit.vfu.cz/statpotr/POTR/Teorie/Predn3/ttest.htm>
- [27] TURČIČOVÁ, Marie. Dvouvýběrové testy: Dvouvýběrový Wilcoxonův test. *Matematická sekce* [online]. Praha: Matematicko-fyzikální fakulta Univerzita Karlova [cit. 2021-3-10]. Dostupné z: https://www2.karlin.mff.cuni.cz/~turcic/Dvouvyberove_testy.pdf
- [28] PAVLÍK, Tomáš a Ladislav DUŠEK. *Biostatistika: Základy korelační analýzy* [online]. Brno: Akademické nakladatelství CERM, 2012 [cit. 2021-5-5]. ISBN 978-80-7204-782-6. Dostupné z: <https://www.matematickabiologie.cz/res/file/ucebnice/pavlik-biostatistika-v2.pdf>
- [29] *Mathsisfun: Correlation* [online]. [cit. 2021-4-6]. Dostupné z: <https://www.mathsisfun.com/data/correlation.html>
- [30] LITSCHMANNOVÁ, Martina. Jednoduchá lineární regrese. *Statistika I., cvičení* [online]. [cit. 2021-5-11]. Dostupné z: https://homel.vsb.cz/~dom033/predmety/statistika/cviceni_priklady/14cRegrese.pdf
- [31] Český statistický úřad: ČSÚ a územně analytické podklady [online]. 2021 [cit. 2021-5-10]. Dostupné z: https://www.czso.cz/csu/czso/csu_a_uzemne_analyticke_podklady
- [32] ROUSSEEUW, Peter J. a Mia HUBERT. Robust statistics for outlier detection. *WIREs Data Mining and Knowledge Discovery* [online]. 2011, 1(1), 73-79 [cit. 2021-5-11]. ISSN 1942-4787. DOI: 10.1002/widm.2 Dostupné z: <https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.2>

SEZNAM POUŽITÝCH ZKRATEK

SPETEP	Strategické partnerství pro enviromentální technologie a produkci energie
EU	Evropská unie
KO	Komunální odpad
OH	Odpadové hospodářství
SKO	Směsný komunální odpad
TIRSMZP719	Prognózování produkce odpadů a stanovování složení komunálního odpadu
ÚPI	Ústav procesního inženýrství

SEZNAM POUŽITÝCH VELIČIN

\mathbf{b}	Matice
b	Prvek matice \mathbf{b}
D_n	Pozorovaná hodnota
$D_n(\alpha)$	Kritická tabelovaná hodnota
$D(X)$	Rozptyl
$\det H_1$	Determinant matice H_1
$E(X)$	Střední hodnota
e	Nekorelované náhodné chyby
$F(x)$	Distribuční funkce
$F_0(X)$	Hypotetickou distribuční funkcí
$F_n(X)$	Výběrová distribuční funkce
$f(x)$	Hustota pravděpodobnosti
\mathbf{g}	Matice
g	Prvek matice \mathbf{g}
H	Hypotéza
\mathbf{H}_1	Matice
k	Stupeň volnosti
m	Hodnost matice
n	Počet členů výběrového souboru (rozsah)
$p(x)$	Pravděpodobnostní funkce
$R(X, Y)$	Pearsonův korelační koeficient
$R(Y)$	Pořadí výběru Y
r	Pearsonův korelační koeficient
r_s	Spearmanův korelační koeficient
S_{min}^*	Reziduální hodnota (metody nejmenších čtverců)
s	Směrodatná odchylka
s^2	Rozptyl
T	Odhad
t	Pozorovaná hodnota
W_α	Kritický obor
\bar{W}_α	Doplňek kritického oboru
W_y	Součet pořadí jednoho výběru
X	Proměnná
\bar{x}	Aritmetický průměr
x	Náhodná veličina
Y	Náhodná veličina
\bar{y}	Aritmetický průměr
y	Pozorovaná hodnota náhodné veličiny Y
Z	Pozorovaná hodnota

α	Hladina významnosti testu
β_0	Bodový odhad regresního koeficientu
β_1	Intervalový odhad regresního koeficientu
ε	Reálné číslo
μ	Směrodatná odchylka
$\sigma(x)$	Směrodatná odchylka
σ^2	Rozptyl
ϑ	Parametr

SEZNAM OBRÁZKŮ

Obr. 1 Příklad grafu pravděpodobnostní funkce diskrétního typu	18
Obr. 2 Příklad grafu distribuční funkce diskrétního typu	18
Obr. 3 Příklad grafu hustoty pravděpodobnosti spojitého typu	19
Obr. 4 Příklad grafu distribuční funkce spojitého typu	19
Obr. 5 Graf hustoty normálního rozdělení	23
Obr. 6 Graf distribuční funkce normálního rozdělení	23
Obr. 7 Příklad hodnot korelačních koeficientů [29]	29
Obr. 8 Průměrné složení odpadu Brna	33
Obr. 9 Průměrné složení odpadu Valašských Klobouků	34
Obr. 10 Průměrné složení odpadu Louky	34
Obr. 11 Schéma znázornění postupu vyhodnocování hypotéz	36
Obr. 12 Schéma výsledků provedených testů	40

SEZNAM TABULEK

Tab. 1 Tabulka závěrečného vyhodnocení hypotéz	26
Tab. 2 Hodnoty datových sad pro jednotlivé frakce.....	32
Tab. 3 Výsledky pro normalitu dat.....	37
Tab. 4 Výsledky Fisherova testu	38
Tab. 5 Výsledky Studentova testu pro stejné/různé rozptyly	39

SEZNAM PŘÍLOH

Příloha 1 Tabulka kritických hodnot pro Kolmogorov-Smirnovův test [24]	59
Příloha 2 Tabulka kvantilů Fisherova rozdělení 1. část [15]	60
Příloha 3 Tabulka kvantilů Fisherova rozdělení 2. část [15]	61
Příloha 4 Tabulka kvantilů Studentova rozdělení [15]	62

PŘÍLOHY

Tab. 6. Kritické hodnoty pro Kolmogorov-Smirnovův test; $D_\alpha(n)$.

n	α					n	α				
	0.2	0.1	0.05	0.02	0.01		0.2	0.1	0.05	0.02	0.01
1	0.900	0.950	0.975	0.990	0.995	31	0.187	0.214	0.238	0.266	0.285
2	0.684	0.776	0.842	0.900	0.929	32	0.184	0.211	0.234	0.262	0.281
3	0.565	0.636	0.708	0.785	0.829	33	0.182	0.208	0.231	0.258	0.277
4	0.493	0.565	0.624	0.689	0.734	34	0.179	0.205	0.227	0.254	0.273
5	0.447	0.509	0.563	0.627	0.669	35	0.177	0.202	0.224	0.251	0.269
6	0.410	0.468	0.519	0.577	0.617	36	0.174	0.199	0.221	0.247	0.265
7	0.381	0.436	0.483	0.538	0.576	37	0.172	0.196	0.218	0.244	0.262
8	0.358	0.410	0.454	0.507	0.542	38	0.170	0.194	0.215	0.241	0.258
9	0.339	0.387	0.430	0.480	0.513	39	0.168	0.191	0.213	0.238	0.255
10	0.323	0.369	0.409	0.457	0.4899	40	0.165	0.189	0.210	0.235	0.252
11	0.308	0.352	0.391	0.438	0.468	41	0.163	0.187	0.208	0.232	0.249
12	0.296	0.338	0.375	0.419	0.449	42	0.162	0.185	0.205	0.229	0.246
13	0.285	0.325	0.361	0.404	0.432	43	0.160	0.183	0.203	0.227	0.243
14	0.275	0.314	0.349	0.390	0.418	44	0.158	0.181	0.201	0.224	0.241
15	0.266	0.304	0.338	0.377	0.404	45	0.156	0.179	0.198	0.222	0.238
16	0.258	0.295	0.327	0.366	0.392	46	0.155	0.177	0.196	0.219	0.235
17	0.250	0.286	0.318	0.355	0.381	47	0.153	0.175	0.194	0.217	0.233
18	0.244	0.279	0.309	0.346	0.371	48	0.151	0.173	0.192	0.214	0.231
19	0.237	0.271	0.301	0.337	0.361	49	0.150	0.171	0.190	0.213	0.228
20	0.232	0.265	0.294	0.329	0.352	50	0.148	0.170	0.188	0.211	0.226
21	0.226	0.259	0.287	0.321	0.344	51	0.147	0.168	0.187	0.209	0.224
22	0.221	0.253	0.281	0.314	0.337	52	0.146	0.166	0.185	0.207	0.222
23	0.216	0.247	0.275	0.307	0.330	53	0.144	0.165	0.183	0.205	0.220
24	0.212	0.242	0.269	0.301	0.323	54	0.143	0.163	0.181	0.203	0.218
25	0.208	0.238	0.264	0.295	0.317	55	0.142	0.162	0.180	0.201	0.216
26	0.204	0.233	0.259	0.290	0.311	56	0.140	0.160	0.178	0.199	0.214
27	0.200	0.229	0.254	0.284	0.305	57	0.139	0.159	0.177	0.198	0.212
28	0.197	0.225	0.250	0.279	0.300	58	0.138	0.158	0.175	0.196	0.210
29	0.193	0.221	0.246	0.275	0.295	59	0.137	0.156	0.174	0.194	0.208
30	0.190	0.218	0.242	0.270	0.290	$n \geq 60$	$\frac{1.07298}{\sqrt{n}}$	$\frac{1.22387}{\sqrt{n}}$	$\frac{1.35810}{\sqrt{n}}$	$\frac{1.51743}{\sqrt{n}}$	$\frac{1.62762}{\sqrt{n}}$

Příloha 1 Tabulka kritických hodnot pro Kolmogorov-Smirnovův test [24]

4. Kvantily F_P Fisherova – Snedecorova rozdělení $F(k_1, k_2)$ pro $P = 0,975$

$k_1 \backslash k_2$	1	2	3	4	5	6	7	8	9	10
1	647,793	799,482	864,151	899,599	921,835	937,114	948,203	956,643	963,279	968,634
2	38,506	39,000	39,166	39,248	39,298	39,331	39,356	39,373	39,387	39,398
3	17,443	16,044	15,439	15,101	14,885	14,735	14,624	14,540	14,473	14,419
4	12,218	10,649	9,979	9,604	9,364	9,197	9,074	8,980	8,905	8,844
5	10,007	8,434	7,764	7,388	7,146	6,978	6,853	6,757	6,681	6,619
6	8,813	7,260	6,599	6,227	5,988	5,820	5,695	5,600	5,523	5,461
7	8,073	6,542	5,890	5,523	5,285	5,119	4,995	4,899	4,823	4,761
8	7,571	6,059	5,416	5,053	4,817	4,652	4,529	4,433	4,357	4,295
9	7,209	5,715	5,078	4,718	4,484	4,320	4,197	4,102	4,026	3,964
10	6,937	5,456	4,826	4,468	4,236	4,072	3,950	3,855	3,779	3,717
11	6,724	5,256	4,630	4,275	4,044	3,881	3,759	3,664	3,588	3,526
12	6,554	5,096	4,474	4,121	3,891	3,728	3,607	3,512	3,436	3,374
13	6,414	4,965	4,347	3,996	3,767	3,604	3,483	3,388	3,312	3,250
14	6,298	4,857	4,242	3,892	3,663	3,501	3,380	3,285	3,209	3,147
15	6,200	4,765	4,153	3,804	3,576	3,415	3,293	3,199	3,123	3,060
16	6,115	4,687	4,077	3,729	3,502	3,341	3,219	3,125	3,049	2,986
17	6,042	4,619	4,011	3,665	3,438	3,277	3,156	3,061	2,985	2,922
18	5,978	4,560	3,954	3,608	3,382	3,221	3,100	3,005	2,929	2,866
19	5,922	4,508	3,903	3,559	3,333	3,172	3,051	2,956	2,880	2,817
20	5,871	4,461	3,859	3,515	3,289	3,128	3,007	2,913	2,837	2,774
21	5,827	4,420	3,819	3,475	3,250	3,090	2,969	2,874	2,798	2,735
22	5,786	4,383	3,783	3,440	3,215	3,055	2,934	2,839	2,763	2,700
23	5,750	4,349	3,750	3,408	3,183	3,023	2,902	2,808	2,731	2,668
24	5,717	4,319	3,721	3,379	3,155	2,995	2,874	2,779	2,703	2,640
25	5,686	4,291	3,694	3,353	3,129	2,969	2,848	2,753	2,677	2,613
26	5,659	4,265	3,670	3,329	3,105	2,945	2,824	2,729	2,653	2,590
27	5,633	4,242	3,647	3,307	3,083	2,923	2,802	2,707	2,631	2,568
28	5,610	4,221	3,626	3,286	3,063	2,903	2,782	2,687	2,611	2,547
29	5,588	4,201	3,607	3,267	3,044	2,884	2,763	2,669	2,592	2,529
30	5,568	4,182	3,589	3,250	3,026	2,867	2,746	2,651	2,575	2,511
35	5,485	4,106	3,517	3,179	2,956	2,796	2,676	2,581	2,504	2,440
40	5,424	4,051	3,463	3,126	2,904	2,744	2,624	2,529	2,452	2,388
45	5,377	4,009	3,422	3,086	2,864	2,705	2,584	2,489	2,412	2,348
50	5,340	3,975	3,390	3,054	2,833	2,674	2,553	2,458	2,381	2,317
55	5,310	3,948	3,364	3,029	2,807	2,648	2,528	2,433	2,355	2,291
60	5,286	3,925	3,343	3,008	2,786	2,627	2,507	2,412	2,334	2,270
70	5,247	3,890	3,309	2,975	2,754	2,595	2,474	2,379	2,302	2,237
80	5,218	3,864	3,284	2,950	2,730	2,571	2,450	2,355	2,277	2,213
90	5,196	3,844	3,265	2,932	2,711	2,552	2,432	2,336	2,259	2,194
100	5,179	3,828	3,250	2,917	2,696	2,537	2,417	2,321	2,244	2,179
120	5,152	3,805	3,227	2,894	2,674	2,515	2,395	2,299	2,222	2,157
150	5,126	3,781	3,204	2,872	2,652	2,494	2,373	2,278	2,200	2,135
250	5,085	3,744	3,169	2,837	2,618	2,459	2,338	2,243	2,165	2,100
500	5,054	3,716	3,142	2,811	2,592	2,434	2,313	2,217	2,139	2,074
∞	5,024	3,689	3,116	2,786	2,566	2,408	2,288	2,192	2,114	2,048

Příloha 2 Tabulka kvantilů Fisherova rozdělení 1. část [15]

**4. Kvantily F_P Fisherova – Snedecorova rozdělení $F(k_1, k_2)$ pro $P = 0,975$
(pokračování)**

$k_1 \backslash k_2$	12	15	20	24	30	40	60	100	250	∞
1	976,725	984,874	993,081	997,272	1001,40	1005,60	1009,79	1013,16	1016,22	1018,26
2	39,415	39,431	39,448	39,457	39,465	39,473	39,481	39,488	39,494	39,498
3	14,337	14,253	14,167	14,124	14,081	14,036	13,992	13,956	13,924	13,902
4	8,751	8,657	8,560	8,511	8,461	8,411	8,360	8,319	8,282	8,257
5	6,525	6,428	6,329	6,278	6,227	6,175	6,123	6,080	6,041	6,015
6	5,366	5,269	5,168	5,117	5,065	5,012	4,959	4,915	4,876	4,849
7	4,666	4,568	4,467	4,415	4,362	4,309	4,254	4,210	4,170	4,142
8	4,200	4,101	3,999	3,947	3,894	3,840	3,784	3,739	3,698	3,670
9	3,868	3,769	3,667	3,614	3,560	3,505	3,449	3,403	3,361	3,333
10	3,621	3,522	3,419	3,365	3,311	3,255	3,198	3,152	3,109	3,080
11	3,430	3,330	3,226	3,173	3,118	3,061	3,004	2,956	2,912	2,883
12	3,277	3,177	3,073	3,019	2,963	2,906	2,848	2,800	2,755	2,725
13	3,153	3,053	2,948	2,893	2,837	2,780	2,720	2,671	2,626	2,595
14	3,050	2,949	2,844	2,789	2,732	2,674	2,614	2,565	2,519	2,487
15	2,963	2,862	2,756	2,701	2,644	2,585	2,524	2,474	2,427	2,395
16	2,889	2,788	2,681	2,625	2,568	2,509	2,447	2,396	2,349	2,316
17	2,825	2,723	2,616	2,560	2,502	2,442	2,380	2,329	2,280	2,247
18	2,769	2,667	2,559	2,503	2,445	2,384	2,321	2,269	2,220	2,187
19	2,720	2,617	2,509	2,452	2,394	2,333	2,270	2,217	2,167	2,133
20	2,676	2,573	2,464	2,408	2,349	2,287	2,223	2,170	2,120	2,085
21	2,637	2,534	2,425	2,368	2,308	2,246	2,182	2,128	2,077	2,042
22	2,602	2,498	2,389	2,332	2,272	2,210	2,145	2,090	2,039	2,003
23	2,570	2,466	2,357	2,299	2,239	2,176	2,111	2,056	2,004	1,968
24	2,541	2,437	2,327	2,269	2,209	2,146	2,080	2,024	1,972	1,935
25	2,515	2,411	2,300	2,242	2,182	2,118	2,052	1,996	1,942	1,906
26	2,491	2,387	2,276	2,217	2,157	2,093	2,026	1,969	1,915	1,878
27	2,469	2,364	2,253	2,195	2,133	2,069	2,002	1,945	1,891	1,853
28	2,448	2,344	2,232	2,174	2,112	2,048	1,980	1,922	1,867	1,829
29	2,430	2,325	2,213	2,154	2,092	2,028	1,959	1,901	1,846	1,807
30	2,412	2,307	2,195	2,136	2,074	2,009	1,940	1,882	1,826	1,787
35	2,341	2,235	2,122	2,062	1,999	1,932	1,861	1,801	1,743	1,702
40	2,288	2,182	2,068	2,007	1,943	1,875	1,803	1,741	1,680	1,637
45	2,248	2,141	2,026	1,965	1,900	1,831	1,757	1,694	1,631	1,586
50	2,216	2,109	1,993	1,931	1,866	1,796	1,721	1,656	1,592	1,545
55	2,190	2,083	1,967	1,904	1,838	1,768	1,692	1,625	1,559	1,511
60	2,169	2,061	1,944	1,882	1,815	1,744	1,667	1,599	1,532	1,482
70	2,136	2,028	1,910	1,847	1,779	1,707	1,628	1,558	1,488	1,436
80	2,111	2,003	1,884	1,820	1,752	1,679	1,599	1,527	1,455	1,400
90	2,092	1,983	1,864	1,800	1,731	1,657	1,576	1,503	1,428	1,371
100	2,077	1,968	1,849	1,784	1,715	1,640	1,558	1,483	1,407	1,347
120	2,055	1,945	1,825	1,760	1,690	1,614	1,530	1,454	1,374	1,310
150	2,032	1,922	1,801	1,736	1,665	1,588	1,502	1,423	1,340	1,271
250	1,997	1,886	1,764	1,697	1,625	1,546	1,457	1,374	1,282	1,201
500	1,971	1,859	1,736	1,669	1,596	1,515	1,423	1,336	1,235	1,137
∞	1,945	1,833	1,708	1,640	1,566	1,484	1,388	1,296	1,183	1,000

Příloha 3 Tabulka kvantilů Fisherova rozdělení 2. část [15]

2. Kvantily t_P Studentova rozdělení $S(k)$

$\begin{matrix} P \\ k \end{matrix}$	0,95	0,975	0,99	0,995	0,999	0,9995
1	6,314	12,706	31,821	63,656	318,289	636,578
2	2,920	4,303	6,965	9,925	22,328	31,600
3	2,353	3,182	4,541	5,841	10,214	12,924
4	2,132	2,776	3,747	4,604	7,173	8,610
5	2,015	2,571	3,365	4,032	5,894	6,869
6	1,943	2,447	3,143	3,707	5,208	5,959
7	1,895	2,365	2,998	3,499	4,785	5,408
8	1,860	2,306	2,896	3,355	4,501	5,041
9	1,833	2,262	2,821	3,250	4,297	4,781
10	1,812	2,228	2,764	3,169	4,144	4,587
11	1,796	2,201	2,718	3,106	4,025	4,437
12	1,782	2,179	2,681	3,055	3,930	4,318
13	1,771	2,160	2,650	3,012	3,852	4,221
14	1,761	2,145	2,624	2,977	3,787	4,140
15	1,753	2,131	2,602	2,947	3,733	4,073
16	1,746	2,120	2,583	2,921	3,686	4,015
17	1,740	2,110	2,567	2,898	3,646	3,965
18	1,734	2,101	2,552	2,878	3,610	3,922
19	1,729	2,093	2,539	2,861	3,579	3,883
20	1,725	2,086	2,528	2,845	3,552	3,850
21	1,721	2,080	2,518	2,831	3,527	3,819
22	1,717	2,074	2,508	2,819	3,505	3,792
23	1,714	2,069	2,500	2,807	3,485	3,768
24	1,711	2,064	2,492	2,797	3,467	3,745
25	1,708	2,060	2,485	2,787	3,450	3,725
26	1,706	2,056	2,479	2,779	3,435	3,707
27	1,703	2,052	2,473	2,771	3,421	3,689
28	1,701	2,048	2,467	2,763	3,408	3,674
29	1,699	2,045	2,462	2,756	3,396	3,660
30	1,697	2,042	2,457	2,750	3,385	3,646
35	1,690	2,030	2,438	2,724	3,340	3,591
40	1,684	2,021	2,423	2,704	3,307	3,551
45	1,679	2,014	2,412	2,690	3,281	3,520
50	1,676	2,009	2,403	2,678	3,261	3,496
60	1,671	2,000	2,390	2,660	3,232	3,460
70	1,667	1,994	2,381	2,648	3,211	3,435
80	1,664	1,990	2,374	2,639	3,195	3,416
90	1,662	1,987	2,368	2,632	3,183	3,402
100	1,660	1,984	2,364	2,626	3,174	3,390
120	1,658	1,980	2,358	2,617	3,160	3,373
140	1,656	1,977	2,353	2,611	3,149	3,361
160	1,654	1,975	2,350	2,607	3,142	3,352
180	1,653	1,973	2,347	2,603	3,136	3,345
200	1,653	1,972	2,345	2,601	3,131	3,340
300	1,650	1,968	2,339	2,592	3,118	3,323
500	1,648	1,965	2,334	2,586	3,107	3,310
1000	1,646	1,962	2,330	2,581	3,098	3,300
∞	1,645	1,960	2,326	2,576	3,090	3,290

Poznámka: Pro $0 \leq P \leq 0,5$ použijeme vztah $t_P = -t_{1-P}$.

Příloha 4 Tabulka kvantilů Studentova rozdělení [15]